DOCUMENT RESUME

ED 263 127 TM 850 585

AUTHOR	Duran, Richard P.; And Others
TITLE	TOEFL from a Communicative Viewpoint on Language Proficiency: A Working Paper.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-85-8; TOEFL-RR-17
PUB DATE	Feb 85
NOTE	187p.
PUB TYPE	Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	College Entrance Examinations; *Communicative Competence (Languages); Content Analysis; Criterion Referenced Tests; *English (Second Language); Higher Education; Item Analysis; *Language Proficiency; *Language Tests; Models; Performance Factors; Second Language Learning; *Skill Analysis; Test Construction; Test Items; Test Reviews
IDENTIFIERS	Interagency Language Roundtable; *Test of English as a Foreign Language

ABSTRACT

The content characteristics of the Test of English as a Foreign Language (TOEFL) are examined from a communicative viewpoint, based on current theory in applied linguistics and language proficiency assessment. The study employed a four-part operational framework. The first component analyzed the communicative characteristics of a language proficiency test based on grammatical, sociolinguistic, and discourse competencies required by test tasks. The second component consists of eight factors which could influence scores--test taking skills, representation of knowledge, language use situation, context and message, artificial restrictions, monitoring factors, affective factors, and strategic factors. The third component consists of judgments of the relevance of item content to academic and social language use. The final component considers the relationship of language tasks in TOEFL items to the criterion-referenced Interagency Language Roundtable scale of language proficiency. Suggestions for improving language proficiency tests are presented, which are based on the principle of eliciting the best performance from examinees. Appendices include a 150-item TOEFL test form; a script for the TOEFL listening comprehension section; taxonomies of communication skills; item analyses based on the communication skill checklists; and skill level descriptions from the Interagency Language Roundtable. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263127

TOEFL

TEST OF ENGLISH AS A FOREIGN LANGUAGE

Research Reports

TOEFL from a Communicative Viewpoint on Language Proficiency: A Working Paper

Richard P. Duran
Michael Canale
Joyce Penfield
Charles W. Stansfield
Judith E. Liskin-Gasparro

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Weidenmiller

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 850 585

The Test of English as a Foreign Language (TOEFL) was developed in 1963 by a National Council on the Testing of English as a Foreign Language, which was formed through the cooperative effort of over thirty organizations, public and private, that were concerned with testing the English proficiency of nonnative speakers of the language applying for admission to institutions in the United States. In 1965, Educational Testing Service (ETS) and the College Board assumed joint responsibility for the program and in 1973 a cooperative arrangement for the operation of the program was entered into by ETS, the College Board, and the Graduate Record Examinations (GRE) Board. The membership of the College Board is composed of schools, colleges, school systems, and educational associations; GRE Board members are associated with graduate education.

ETS administers the TOEFL program under the general direction of a Policy Council that was established by, and is affiliated with, the sponsoring organizations. Members of the Policy Council represent the College Board and the GRE Board and such institutions and agencies as graduate schools of business, junior and community colleges, nonprofit educational exchange agencies, and agencies of the United States government.

A continuing program of research related to TOEFL is carried out under the direction of the TOEFL Research Committee. Its six members include representatives of the Policy Council, the TOEFL Committee of Examiners, and distinguished English-as-a-second-language specialists from the academic community. Currently the committee meets twice yearly to review and approve proposals for test-related research and to set guidelines for the entire scope of the TOEFL research program. Members of the Research Committee serve three-year terms at the invitation of the Policy Council; the chair of the committee serves on the Policy Council.

Because the studies are specific to the test and the testing program, most of the actual research is conducted by ETS staff rather than by outside researchers. However, many projects require the cooperation of other institutions, particularly those with programs in the teaching of English as a foreign or second language. Representatives of such programs who are interested in participating in or conducting TOEFL-related research are invited to contact the TOEFL program office. Local research may sometimes require access to TOEFL data. In such cases, the program may provide this data following approval by the Research Committee. All TOEFL research projects must undergo appropriate ETS review to ascertain that the confidentiality of data will be protected.

Current (1984-85) members of the TOEFL Research Committee include the following:

Henry F. Holtzclaw, Jr. (chair)	University of Nebraska
Kathleen M. Bailey	Monterrey Institute of International Studies
Alison d' Anglejan-Chatillon	University of Montreal
H. Douglas Brown	San Francisco State University
Russell N. Campbell	University of California at Los Angeles
John Haskell	Temple University-Japan

TOEFL from a Communicative Viewpoint on Language Proficiency: A Working Paper

Richard P. Duran

Michael Canale

Joyce Penfield

Charles W. Stansfield

Judith E. Liskin-Gasparro

Educational Testing Service
Princeton, New Jersey

RR 85-8

The opinions and views expressed in this report do not necessarily reflect the positions or policies of Educational Testing Service or the Test of English as a Foreign Language Program.

Abstract

This report examines the content characteristics of the Test of English as Foreign Language (TOEFL) from a communicative viewpoint based on current theory in applied linguistics and language proficiency assessment. After a review of relevant literature, the authors developed and applied a four-part operational framework for analyzing the communicative characteristics of a language proficiency test. The first component of this framework consists of the grammatical, sociolinguistic, and discourse competencies required by test tasks. The second component consists of eight factors that could influence test performance. The third component consists of judgements of the relevance of the content of test items to academic and social language use. The final component relates the langauge and language tasks appearing in TOEFL items and sections to a criterion-referenced scale of language proficiency. In this case, the Interagency Language Roundtable scale was used. Finally, the report discusses test design features that might improve the quality of language proficiency tests.

Copyright © 1985 by Educational Testing Service. All rights reserved.

Unauthorized reproduction in whole or in part is prohibited.

TOEFL is a trademark of Educational Testing Service, registered in the U.S.A. and in many other countries.

Preface

This report examines the content characteristics of the Test of English as a Foreign Language (TOEFL) from a communicative perspective based on current research in the area of applied linguistics and language proficiency assessment. The method of analysis used in the study represents one among a number of possible approaches. Communicative competence is a relatively recent perspective on language use. This approach emphasizes the effects of situational context on the use and meaning of language forms. Such an approach would have been considered inappropriate two decades ago, when the emphasis of linguistic analyses was on discrete units, generally considered apart from context. Only in time will it be possible to assess the utility of a situational perspective on language to language assessment practice.

The approach adopted is not intended to address psychometric issues that arise in evaluating the construct and criterion related validity of the TOEFL test; these important issues deserve attention in the context of a report with that specific purpose. The goal of this work is to describe the content characteristics of TOEFL items and sections in terms of a framework for communicative competence and in terms of related factors that can affect performance on language proficiency tests. As noted in the introduction to the report, the framework used to describe TOEFL is exploratory in nature and the detailed findings reported should be viewed as such. A great deal of reflection and judgment is necessary in any content validity study and, thus, in applying the framework to describing TOEFL. As a result, different investigators might vary in the detailed findings that would emerge from their interpretation and application of the framework to the description of TOEFL.

It is important to note also that the report rarely takes into account the actual process by which the TOEFL test is developed and produced. The absence of the viewpoint of test development professionals has allowed the investigators to render an independent description of TOEFL. While such independence has its virtue in terms of the intellectual process of conducting the work described, there are also shortcomings that arise from this procedure. Professionals responsible for the development and production of TOEFL tests along with the actual TOEFL test specifications represent important resources for a fuller understanding of the content characteristics of TOEFL. This should be kept in mind when reading the report.

R.P.D.
C.W.S.

Foreword

The research described in this report required the collaboration of a number of investigators. The research team involved specialists in several overlapping content areas: foreign language teaching and proficiency assessment, applied linguistics and sociolinguistics, and psycholinguistics. Michael Canale of OISE and the University of Toronto authored Section II of the report, concerning advances in language proficiency assessment research and a theoretical framework for evaluating the characteristics of TOEFL. He also authored portions of Section V concerning possible improvements in proficiency assessment instruments and contributed a preliminary checklist of communicative skills, based on Munby (1978) and Canale and Swain's (1980) previous research. Joyce Penfield of Rutgers University, assisted by Richard Duran, authored Sections III and IV of the report with contributions by Charles Stansfield and Judith Liskin-Gasparro. These discuss the use of an expanded skills checklist to describe TOEFL, and the results of applying the checklist to describe a TOEFL instrument. Scoring of the sample TOEFL test was done by Penfield and Richard Duran. Charles Stansfield and Judith Liskin-Gasparro of Educational Testing Service (ETS) authored the portions of Section IV discussing use of the ILR scale to describe the proficiency levels of TOEFL test sections and items. Richard Duran of ETS directed the project and authored the remaining portions of the report. All of the collaborators and their assistants contributed to the final editing of the report, which was coordinated by Richard Duran and Charles Stansfield.

Thanks are extended to the TOEFL Research Committee, the TOEFL Committee of Examiners, and the TOEFL Program for feedback on earlier versions of this report; the authors have made discretionary revisions of the report based on this feedback. Thanks are also extended to E. Leta Davis, Michaele Mikovsky, and Peggy Thorpe for their coding and transcription of data concerning the TOEFL test analyzed. Davis also assisted in coordinating aspects of the final report development. Special thanks go to Jessie Cryer, Thelma Benton, Faith Thompson, and Marian Helms for their typing and manuscript preparation. Finally, thanks are extended to Nancy Parr for editing and proofreading the report manuscript.

Table of Contents

		Page
I.	Introduction	1
	1. Background to the Report	1
	2. TOEFL Sections and Formats	1
	3. Overview of Project Activities and the Report	4
II.	Investigating the Language Skills Required by TOEFL	5
	1. Recent Research in Language Testing	5
	2. Construct Validity Concerns	6
	3. A Framework for Analyzing TOEFL	10
III.	Procedures for Describing the TOEFL	12
	1. Overview	12
	2. Procedures for Conducting the Competence Domain Description of a Sample TOEFL Test	12
	3. Description of the Competence Skills Checklist	13
	3.1 Grammatical Competence (Section A)	13
	3.2 Sociolinguistic Competence (Section B)	18
	3.3 Discourse Competence (Section C)	19
	3.4 Coding of TOEFL Items	20
	4. Procedure for Analyzing Performance Features	21
	5. Procedure for Evaluating the Relevance of TOEFL Item Content to Academic and Social Language Use Contexts	21
	6. Procedures for Preliminary Description of Proficiency Level Represented by TOEFL Items	22
IV.	Findings	24
	1. Plan of Discussion	24
	2. Description of TOEFL Based on the Checklist of Communicative Skills	24

Page

2.1 Summary of Findings for the TOEFL Test as a Whole . 24

2.2 Summary of Findings for Listening Comprehension, Section 1 29

2.2.1 Section 1: Statements, Part A. 32

2.2.2 Section 1: Dialogues, Part B 33

2.2.3 Section 1: Minitalks/Extended Conversation, Part C. 34

2.3 Summary of Findings for Structure and Written Expression, Section 2 35

2.4 Summary of the Findings for Vocabulary and Reading Comprehension, Section 3. 36

2.4.1 Section 3: Vocabulary. 36

2.4.2 Section 3: Reading Comprehension 38

2.5 Concluding Comment on the Checklist Findings. . . . 40

3. Findings Regarding Test Performance Factors Cited by Bachman and Palmer. 40

3.1 Overview of Issues. 40

3.2 Psychophysiological Skills. 40

3.3 Representation of Knowledge About Language and Monitoring Factors. 41

3.4 Language Use Situations 42

3.5 Context and Message 42

3.6 Artificial Restrictions 44

3.7 Affective Factors 45

3.8 Strategic Factors 45

4. Findings Regarding the Relevance of TOEFL Item Content for Academic and Social Language Use. 45

Page

4.1 Rating Agreement. 45

4.2 Ratings of the Characteristics of Section 1 Items . 46

4.3 Ratings of the Characteristics of Section 2 Items . 47

4.4 Ratings of the Characteristics of Section 3 Items . 47

5. ILR Scale Description of TOEFL Proficiency Level. 47

5.1 Section 1: Listening Comprehension 48

5.2 Section 2: Structure and Written Expression. . . . 53

5.3 Section 3: Reading Comprehension and Vocabulary. . 54

5.4 Conclusion 56

V. Discussion and Directions for Item and Test Development. 58

1. Overview of This Section. 58

2. Conclusions Regarding TOEFL 58

3. Suggestions for Improving Language Testing. 61

3.1 Two Guiding Principles. 61

3.2 Four Test Design Features 62

4. Implications for TOEFL Program Activities 64

4.1 Study of the TOEFL Test Specifications and Continued Study of the Content Characteristics of TOEFL Items 64

4.2 Development Research on Thematic Presentation of Items and on New Item Types 64

4.3 New Approaches to Assess Communicative Proficiency. 65

4.4 Technology, Innovative Measurement and Adaptive Testing 65

4.5 Validity Research 66

4.6 Staying Abreast with the Assessment Field 67

Page

References 68

Appendix A: Sample TOEFL Test

Appendix B: General Outline of Communication Skills

Appendix C: Checklist of Competence Areas for Evaluating the TOEFL

Appendix D: Summary of Checklist Results

Appendix E: Interagency Language Roundtable Skill Level Descriptions

I. INTRODUCTION

1. Background to the Report

Advances in foreign language teaching and language assessment research over the past decade have suggested that language proficiency testing theory and practice might benefit from an examination of issues related to communicative competence. This report summarizes findings of a project that investigated the relevance of some of these issues to the Test of English as a Foreign Language based on a theoretical model of communicative competence and the validity of language tests. The orientation toward communication and language use adopted in this report is representative of but one approach that is current in the field of foreign language teaching and testing. It is important to note at the outset that the TOEFL test is intended to evaluate only certain aspects of the English proficiency of persons whose native language is not English. A communicative description of TOEFL should be expected to be sensitive to what the test currently measures as well as to aspects of proficiency that it does not measure. The examination of the TOEFL test based on the communicative approach to language assessment advocated in this report is exploratory. It is based on some widely cited and discussed perspectives on the communicative properties of tests, but the approach taken should not be considered to provide the definitive account of the communicative properties of the TOEFL. Before overviewing the approach of the present project, it will be helpful to discuss the purpose and nature of TOEFL. Special attention will be given to describing the content and format of TOEFL test items since these are of central concern in the present study.

The TOEFL test is intended to evaluate certain aspects of the English language proficiency of persons whose native language is not English. The test is used widely by colleges and universities in the United States and Canada in reviewing the admissions qualifications of incoming foreign students whose native language is not English. In addition, it is used by other organizations and by some agencies of the U.S. government in assessing the English language qualifications of nonnative persons.

Use of the current three-section TOEFL began in 1976; up to that time TOEFL was administered as a five-section test. The total TOEFL score, computed from scores earned on each of three sections, range from 227 to 677. Further information on the use of TOEFL scores is provided in the Test and Score Manual (TOEFL, 1983); attention is focused here on the nature of the three TOEFL sections and the characteristics of their item types.

2. TOEFL Sections and Item Formats

The three sections of TOEFL are Section 1, Listening Comprehension; Section 2, Structure and Written Expression; and Section 3, Reading Comprehension and Vocabulary. The following descriptions of the sections call attention to the formats of items and to the skills tested by items. Some of the skills are directly relevant to components of communicative competence identified later in this report.

Listening Comprehension Section. This section is intended to measure the examinee's ability to understand English as it is spoken in the United States. The three item types in the 50-item section are statements (20 items), dialogues (15 items), and minitalks/extended conversations (15 items).

Statement items present examinees with short recorded statements by individual speakers. After hearing a statement, examinees select one of four response options (given in the test book) that is synonymous with the original statement. All of the response options for an item are either full sentences or elliptical responses. Statement items are designed to test examinees' ability to comprehend a spoken sentence based on its syntactic structure, word meaning and word usage, phonological and suprasegmental characteristics, and inferences that follow from the meaning of a sentence. Incorrect response options reflect inappropriate recognition of words, sentence structures, and speech act types that might result if the sound characteristics of words and sentence elements are not properly recognized. Incorrect responses also might arise if examinees do comprehend written answer options--this, indeed, is a possibility for all sections and item types found on the TOEFL.

Dialogue items present examinees with a recorded conversation between two speakers. In most instances, each participant in the conversation speaks once. Following the exchange a third recorded speaker asks the examinees a question about what was said and the examinees select one of four response options given in the test book. The oral comprehension demands of dialogue items are similar to those of statement items, but they also require that examinees understand the interrelationship between turns in a conversational segment. The questions posed by the third speaker are intended to require comprehension of all conversational turns, rather than one turn alone. Questions probe examinees' recognition of the syntax of the speakers' utterances, the content of what is said, and the speech acts and interrelationships among speech acts represented by the speakers' utterances. Some questions probe examinees' knowledge of word meanings, idioms, and paralinguistic cues in the context of conversation. Finally, some questions test examinees' ability to draw inferences about the characteristics of the speakers or the situations represented by the conversations.

A minitalk/extended conversation item stem consists of a recording of an extended utterance by one speaker or a recording of a conversation involving several turns between two speakers; an additional speaker then asks a series of questions. Examinees respond to each question by selecting one of four response options, printed in their test books, that are in the form of phrases or full sentences. These questions stress examinees' understanding of the situation depicted by a talk or conversation, the roles of the speakers, the intentions of the speakers, the topics of discussion, and the content of what was said. According to the TOEFL test specifications, minitalk/extended conversation items involve situations and topics representative of language use in academic settings or in the everyday life of college students.

Structure and Written Expression Section. This section consists of 15 structure and 25 written expression items, all of which are printed in the test books. Each structure item consists of a sentence with an omitted word.

Examinees are required to select one of four response options consisting of a word or phrase that will complete the sentence. Only one response option is grammatically correct given the structure and content of the stimulus sentence. Items are designed to systematically probe examinees' control of syntactic features appropriate to various clause and phrase types, negative constructions, comparative forms, word ordering within sentences, and statements of parallel relationships at the phrase and clause level.

Written expression items consist of sentences, each with four underlined words corresponding to response options. Examinees are required to select the one response option corresponding to a word that is grammatically inappropriate in the context of a sentence. Grammatical points tested by these items involve agreement in the syntactic and semantic characteristics of function and content words, appropriate usage of function words, appropriate word ordering, appropriate diction and idiomatic usage, appropriate and complete clause or phrase structure, and maintenance of parallel forms. Errors in sentences reflect inappropriate omission or inappropriate inclusion of sentence elements related to the foregoing grammatical features.

Reading Comprehension and Vocabulary Section. This section is designed to measure examinees' ability to understand various kinds of reading material and the meaning and use of words. The section includes 30 vocabulary items followed by 30 reading comprehension items; all are printed in the test book. Each vocabulary item consists of a single sentence (usually, but not always, of a complex form) with one word underlined. From the four response options, examinees are to select the one that is synonymous in meaning to the underlined word. The occurrence of grammatical category and the difficulty of vocabulary is controlled for in both the item and the response options. According to the TOEFL test specifications, sentences are representative of materials that students might encounter in books, magazines, articles, newspapers, and other printed matter.

Each reading comprehension item consists of a stimulus passage followed by a series of items pertaining to the main idea of the passage, supporting ideas or facts related to the passage, inferences based on passage information, the attitude or tone of the passage, or the application of passage information to other information. Examinees answer each question by selecting one of four answer options, which are sometimes phrases and sometimes full sentences. Some item stems are in the form of incomplete statements that must be completed with one of four response options that are phrases or clauses. Regardless of format, there is only one correct response for each reading comprehension item.

Passages used in the reading comprehension section are 100 to 350 words long and are followed by 3 to 8 questions. The specifications for the design of passages and items controls for the vocabulary characteristics, rhetorical organization of passages, and content characteristics of passages. According to the test specifications, the passages used in the reading comprehension section are representative of materials that students might encounter in school work or in situations related to college life.

A sample TOEFL test is presented in Appendix A. This test, identified as Form 3FATF5, was administered on May 14, 1983, and is the test that is described in detail in this report.

3. Overview of Project Activities and the Report

The central goal of the project was to describe the TOEFL test in light of an approach based on recent theory and research on language proficiency assessment. The outcome of this description included some suggestions for further research on the content and construct validity of TOEFL, as well as for item and test development research. The underlying motivation was an awareness of a significant shift in language teaching pedagogy and language assessement theory that has stressed extended knowledge of how to use language in the everyday world. Accordingly, the first set of activities was to draw on the expertise of a language assessment researcher attuned to leading developments in the field who could summarize current developments in the art and science of language testing and provide a useful framework from which to view significant issues that needed treatment in describing the range of language proficiency skills assessed and not assessed by the test. The summary of important trends in the field and an overview of a framework for describing TOEFL and other proficiency tests are given in Section II.

The second major set of activities was to create procedures for utilizing the framework that was developed and then proceed to study a sample TOEFL instrument in light of the framework. We viewed the effort to operationalize and apply the framework as exploratory. To the best of our knowledge, no similar attempt has been made to describe TOEFL or any other language proficiency instrument in this manner. We realize we have much to learn and that our modest efforts were only a beginning toward a more adequate description of the content validity of language proficiency tests in general. Development of the procedures for describing the TOEFL test based on the framework is described in Section III. The resulting description of the sample TOEFL is in Section IV.

A final objective of our work was to suggest ways in which language proficiency tests might be improved with regard to both usefulness and validity given the findings of our review of TOEFL. While our concern lay most with TOEFL and how its strengths and limitations might be evaluated, we were also concerned with presenting a position on improvements that would be of value to the field of language testing as a whole. Further, it seemed important to suggest concrete steps that might be undertaken to devise new instruments for assessing extended language proficiency. Section IV of the report is devoted to these matters, and also summarizes the findings of our research on TOEFL.

II. INVESTIGATING THE LANGUAGE SKILLS REQUIRED BY TOEFL

1. Recent Research in Language Testing

In recent state-of-the-art reports, Spolsky (1979 and elsewhere) has distinguished three major approaches to language testing. In brief, the traditional approach has viewed language testing as an art best practiced by the language teacher, and has focused largely on what should be tested rather than on how testing should be done. The psychometric-structuralist approach views language testing as a science best practiced by measurement scientists (psychometricians) working with language scientists (linguists); the focus is on both how to test and what to test. Finally, the integrative-sociolinguistic approach is in many ways a synthesis of the other two. In this approach, language testing is viewed as both an art and a science, requiring the collective expertise of language teachers, linguists, psychometricians, sociologists, and anthropologists; the focus is on what is involved in authentic language use (sociolinguistic perspectives) and on how to measure this use through language tests (integrative techniques).

A concern expressed frequently by proponents of this third approach is that many major language tests, under the influence of the psychometric-structuralist approach, may be more objective, reliable, and administratively efficient than they are valid, acceptable, and relevant to the needs of examinees and teachers (e.g., Oller & Spolsky, 1979). This concern is also expressed about psychological and educational testing in general. For example, Messick (1981) has argued that the two most important components in psychological and educational testing are construct validity (roughly, how well we can demonstrate what we are or should be testing) and the potential social consequences of test results. Moreover, Shoemaker (1980) has assembled the sobering viewpoints of such testing experts as Buros, Thorndike, and Lumsden on the extent to which achievement testing has progressed during this century; the view of Buros (1977), reported below, is typical.

> Except for the tremendous advances in electronic scoring, analysis, and reporting of test results, we don't have a great deal to show for fifty years of work. Essentially, achievement tests are being constructed today in the same way they were fifty years ago--the major changes being the use of more sophisticated statistical procedures for doing what we did then--mistakes and all.... In fact, some of today's tests may even be poorer, because of the restrictions imposed by machine scoring (p. 10).

Reflecting these concerns, much recent research in language testing has addressed two issues: the construct validity of language tests and the relevance of language tests to the needs of students and teachers. We turn to some examples of this research on construct validity; the particular studies mentioned are not the only noteworthy research, but are intended to serve as convenient examples.

2. Construct Validity Concerns

Two main lines of research on constuct validity are summarized here: one dealing with theoretical frameworks of language proficiency, the other with the nature of performance on language tests.

Theoretical frameworks for language proficiency. Inspired by the work of such researchers as Vygotsky (1962), Chomsky (1965), Britton (1970), Gumperz (1971), Halliday (1973), Hymes (1974), Bruner (1975), and others, much research on language proficiency has attempted to explain three widespread and nontrivial findings in recent work on language proficiency testing (see Canale 1983a, 1983b, and 1984; and Cummins 1981 and 1983 for references).

a. Certain individuals (often members of language minority groups) have been misclassified as having language disorders and "linguistic deficits," that is as lacking, in a biological/maturational sense, basic language proficiency.

b. Certain individuals who have studied a second language in a formal classroom setting, and who perform well on academically oriented second language tests, do not perform as well on tests requiring use of the second language for authentic communication outside such classroom settings.

c. Certain individuals who perform well on tests requiring authentic communication in the second language do not perform as well on academically oriented second language tests requiring autonomous language use--such as in organizing one's ideas or solving a verbal math problem.

To address such findings adequately and to respond to other shortcomings of previous frameworks for language proficiency, it has been suggested that a theoretical framework with the following three general features is needed: first, basic, communicative, and autonomous language proficiencies should be distinguished; second, the types of knowledge and skills (competence areas) involved in each dimension of language proficiency should be identified; and third, the linguistic and other cognitive demands of a given language task should be considered separately (cf., Canale, 1983a, for a discussion). Following is a brief outline of both the minimal range of language-related competence that could be considered and the manner in which it might contribute to the three dimensions of language proficiency.

As a preliminary range of language competence areas, one might consider those discussed by Canale and Swain (1980). Although proposed originally for describing only communicative language use, these competence areas may represent core components of overall language proficiency and hence may be useful for understanding other uses of language as well. Since these areas are discussed in detail (with references to other work) by Canale and Swain (1980) and Canale (1983b), and in the next part of this report, what follows is only an outline of their essential features.

1. Grammatical competence: mastery of the language code (e.g., vocabulary and rules of word formation, sentence formation, literal meaning, pronunciation, and spelling).

2. Sociolinguistic competence: mastery of appropriate use (production and comprehension) of language in different sociolinguistic contexts, with emphasis on appropriateness of (a) meanings (e.g., topics, attitudes, functions) and (b) forms (e.g., register, formulaic expressions).

3. Discourse competence: mastery of how to combine meanings and forms to achieve unified text in different genres—such as a casual conversation, an argumentative essay, or a business letter--by using both (a) cohesion devices to relate forms (e.g., use of pronouns, transition expressions, parallel structures) and (b) coherence principles to organize meanings (e.g., concerning the development, consistency, balance, and completeness of ideas).

4. Strategic competence: mastery of verbal and nonverbal strategies both (a) to compensate for breakdowns in communication due to insufficient competence or to performance limitations (e.g., use of paraphrase, dictionaries) and (b) to enhance the rhetorical effect of language (e.g., use of slow and soft speech, use of literary devices).

With this range of competencies in mind, consider the three dimensions of language proficiency proposed below.

Basic language proficiency. This dimension, while not critical to the substance of this report, is mentioned here for the sake of completeness. It is concerned with the biological universals required for normal language development and use. Of concern, then, are not only universals of grammar that underlie grammatical competence but also sociolinguistic universals, discourse universals, universal strategies, and perceptual/processing universals. It is assumed that such universals interact with general cognitive development to set the biological upper and lower limits on possible uses, messages, and forms of language. A normal individual's actual communicative and autonomous proficiencies are thus presumably within these limits and are influenced mainly by socialization and, to perhaps a relatively lesser extent, by individual differences in personality, intelligence, aptitudes, learning style, motivation, and personal experiences.

Communicative language proficiency. The focus here is on social, interpersonal (or other-directed) uses of language through spoken or written channels, similar to what Vygotsky (1962) and others have referred to as social speech. It is assumed that communication is primarily a form of social interaction, in which emphasis is normally less on grammatical forms and literal meaning than on participants and what they are trying to do through language--i.e., on the social meaning of utterances. Such social meaning is qualified by contextual variables, such as roles of participants, setting, goals, and norms of interaction; authentic communication thus requires continuous evaluation, interpretation, and negotiation of various levels of information. Such contextual variables may serve to simplify communication

(by providing clues to meaning) or complicate it (e.g., by imposing language-specific appropriate conditions). Although communicative language use normally involves grammatical, sociolinguistic, discourse, and strategic competencies (as identified above), the focus and greatest demand may be on sociolinguistic and strategic knowledge and skills, less so on discourse ones (e.g., cohesion devices and genre structures), and least so on grammatical competence As such, the degree of exposure to and use of sociolinguistic rules, communication strategies, and interactional discourse features--that is, degree of socialization and acculturation with respect to a given language community--may be especially important in determining the range of communicative contexts and functions that an individual can and is willing to handle through that language.

Autonomous language proficiency. This dimension involves less directly social and more intrapersonal, representational (or self-directed) uses of language, such as problem solving, organizing one's thoughts, verbal play, poetry, and personal writing. It is similar to what Vygotsky (1962) and others have referred to as inner speech. Focus is less on social meaning than on personal and literal meanings; hence, contextual variables do not seem to qualify (simplify or complicate) information as much as do the language code, logical relationships among propositions, and idiosyncratic personal interpretations of these. Though immediate sociolinguistic context may be rich, it is not necessarily in focus in autonomous language uses (for example, in counting one's change at the local grocery store). The main language competencies involved would seem to be grammatical (especially vocabulary and rules for deriving literal meaning) and discourse (especially coherence principles), with less demand on communicative strategies and the least on sociolinguistic competence. Again, however, degree of socialization--e.g., degree of exposure to and acceptance of various autonomous tasks in a given language--may be viewed as a valuable index of the range of such tasks that can be performed by an individual through that same language without undue affective, linguistic, and general cognitive difficulties.

To summarize, the relationships among these three dimensions of language proficiency seem to be as follows. Basic language proficiency is comprised of those language-related biological universals that are required for communicative and autonomous language development and use. Communicative and autonomous proficiencies seem to differ in that sociolinguistic and strategic competencies may receive more emphasis in the former whereas grammatical and discourse competencies may be more in focus in the latter. In this view, it follows that one cannot adequately develop or test communicative proficiency through autonomous tasks or vice versa.

The nature of performance on language tests. Just as lack of understanding of what we are trying to measure limits the construct validity of language tests, so does our lack of understanding of how different test methods (or formats) influence performance on such tests. Careful studies have been carried out to investigate just how examinees' attitudes, perceptions, and training with respect both to formal testing and to a particular test method (e.g., multiple-choice) can affect a test score (e.g., Deyhle, 1983; Scott & Madsen, 1983); compelling arguments have been made for paying

close attention to the quality of performance elicited by various test methods (e.g., Jones, 1983). The research reviewed here is that of Bachman and Palmer (forthcoming). They examine eight factors influencing language performance that may, in addition to an examinee's language proficiency, be reflected in a language test score. The important point they stress is that the type of performance elicited through certain test methods may be qualitatively different from that involved in authentic language use; to the extent that test performance does differ importantly and unpredictably from authentic performance, it may in some instances prove difficult to use a test score to draw firm conclusions about an examinee's true ability to use language.

Before outlining the eight factors cited by Bachman and Palmer, it is important to note that the range of authentic language to which learners of English as a foreign language have been exposed to is likely to be limited. Therefore, in some sense, it is not appropriate to expect that an English proficiency test such as TOEFL, used for college and university admission of foreign students, would be able to cover the widest range of authentic language use that might be judged to occur in the life of college students in the United States.

The eight factors identified by Bachman and Palmer are outlined below; then an illustration is provided to indicate how these factors can be useful in test analysis and development.

Psychophysiological skills. This factor reflects the extent to which sensory skills are called for in receptive language tasks (auditory or visual) and to which neuromuscular skills are required in production tasks (articulatory or digital).

Representation of knowledge. Here the focus is on the extent to which the knowledge required in a language task is consciously or subconsciously stored, on the one hand, and stored as a prefabricated pattern or as a rule, on the other hand.

Language use situations. Such situations may differ along two dimensions: first, the extent to which the situation is reflexive (focused on language itself) or transitive (focused on doing something with language); and second, the extent to which the situation is reciprocal (providing immediate feedback and interaction) or nonreciprocal (providing no immediate feedback nor interaction).

Context and message. Four features are distinguished here: the amount of context that must be attended to (i.e., the ratio of familiar to new information); the distribution of the message (compact or diffuse over time and space); type of message (abstract or concrete information); and control of message (whether or not the information must be consciously dealt with).

Artificial restrictions. Language performance may be limited in contrived ways at several levels: organization of discourse (e.g., changing topic from one test item to another); message (e.g., no illocutionary force);

forms (e.g., contrived use of synonymy in a text to satisfy readability formula); participants (e.g., the social and personal roles assumed); and channels (e.g., a minilecture without any visual aids).

Monitoring factors. Based on Krashen's (1982) notion of the language use strategy of consciously editing or monitoring one's language production, this set of factors includes the time available to perform a task and the extent to which the task encourages attention to form and to the use of explicit, conscious language knowledge.

Affective factors. Four elements are signaled here: preference for field dependence (overall image) or independence (isolated bits of image); inhibition (e.g., in role play or oral interaction); tolerance for ambiguity (e.g., detailed versus general task description); and motivation (e.g., to sound or act like a native of the target language).

Strategic factors. This set of factors deals with the extent to which one is able to resolve various language use problems (e.g., use of paraphrase, gestures, basic sentence patterns, requests for repetition) or to enhance one's performance (e.g., use of literary devices in writing or of rhetorical devices in speaking).

Like the research on theoretical frameworks for language proficiency summarized in the first part of this section, research on the factors outlined in relation to performance on language tests bears on the issue of construct validity in language testing. Together these two sets of studies may provide a coherent basis for carrying out research both on what current language tests may actually be measuring and on how to improve language testing. The following discussion illustrates these possible applications in the context of TOEFL.

3. A Framework for Analyzing TOEFL

The framework developed by the project staff for overviewing the TOEFL test attempted to operationalize the concerns raised in Section I on the construct validity of language tests and their relevance to learners in North American postsecondary educational institutions. This framework has four components that were developed independently from the actual TOEFL test specifications. Consistent with the original research proposal for the work described, a modification of Munby's (1978) classification of skills was used for describing skills required by TOEFL. However, the framework also addressed other issues related to language performance on TOEFL. The framework can be summarized as follows.

1. A domain description of competence areas. The purpose of this description is to identify in sufficient but manageable detail the main knowledge and skills provided by TOEFL in the areas of grammatical, sociolinguistic, discourse, and strategic competencies. It has been developed independently of any consideration of TOEFL; rather, it reflects the domain descriptions presented by Munby (1978) and Canale (1983a, based on Canale &

Swain, 1980). In the present context, this outline was adjusted and developed to reflect important features in the teaching of English as a foreign language.

2. Language performance features. Based on the work of Bachman and Palmer (forthcoming), this component attempts to identify the main features of language performance elicited by TOEFL items. The purpose of this component is to allow preliminary characterization of TOEFL item formats in terms of the extent to which each performance feature is or is not present across various item types and sections of the test.

3. Relevance for academic and social language uses. The purpose of this component is to allow independent judges to characterize, in a scalar manner, the content of TOEFL items, i.e., the relevance of topics, communicative functions, discourse genres, and sociocultural contexts to academic language use, on the one hand, and to social language use, on the other.

4. Minimum mastery level. The ultimate purpose here is to allow independent judges to characterize the difficulty of TOEFL sections and items with reference to criterion performance descriptions. A first step in investigating what level of proficiency is sufficient to perform accurately on TOEFL entails a description of TOEFL items in terms of an appropriate performance scale.

The next part of this report decribes the implementation of this framework in describing the TOEFL test in detail. Before proceeding, it is useful to note that the four areas designated by the framework are important for the evaluation of any proficiency test, not just TOEFL. The generality of the approach is important; it emphasizes the extended validity of a proficiency test. All four areas cited are important to examinees taking a proficiency test and to proficiency test consumers; all four have significant bearings on the legitimacy of test score interpretation and the use of this information for valid purposes. This viewpoint contrasts with other approaches to proficiency testing, which have so often ignored the larger range of multiple links that must be established and evaluated between performance on an instrument and performance in the real world.

A second matter that deserves attention is the exploratory character of applying the framework. The areas designated in the framework are of clear significance, but the process of implementing their interpretation in describing a proficiency test is an experiment. This fact should motivate the reader to question how refinement and implementation of the framework might be improved in future contexts. What follows is an honest, and occasionally critical, portrayal of the successes and difficulties encountered in applying this framework to describing the TOEFL test.

III. PROCEDURES FOR DESCRIBING THE TOEFL

1. Overview

The TOEFL instrument found in Appendix A was described in terms of the four framework components discussed in the previous sections: (1) a domain description of competence areas based on a communicative skills checklist; (2) an analysis of the language performance features of TOEFL sections and item types; (3) an evaluation of the relevance of TOEFL section and item contents to language use in criterion settings; and (4) a preliminary description of the minimum mastery levels represented by TOEFL test sections and item types. The design of procedures to operationalize each of the four framework components is discussed in this portion of the report.

2. Procedures for Conducting the Competence Domain Description of a Sample TOEFL Test

One of the major goals of this study was to conduct a detailed examination of the various communicative skills needed to fulfill language tasks on the TOEFL test. To achieve this goal, a taxonomic list of identifiable skills entering into the communicative process was created. The domain description of competence discussed in the previous section and at length in Canale and Swain (1980) and Canale (1983b) served as the basic analytical perspective of communicative competence. From three of the four components of communicative competence presented in this perspective--grammatical, sociolinguistic, and discourse--an operationally oriented list of skills was suggested by Canale (Appendix B). Strategic competence was excluded as an area since examinees are not required to produce language as part of the TOEFL examination. The checklist in Appendix B was subjected to numerous revisions by Penfield and Duran throughout the project. This resulted in a fairly detailed list that was, in turn, modified later in the project in order to be more generic. Appendix C presents this generic skills checklist, and Appendix D summarizes data resulting from application of this checklist to the sample TOEFL test in Appendix A.

Coding an operational instrument for communicative competence that could be applied to natural communication or simulations of it was a monumental task that deserves some discussion here. At each stage in the process of designing successive versions of the skills checklist, the inadequacies of a taxonomic approach to communication became more apparent. It is clear that there is a strong interaction among elements occurring across areas of the checklist and that the taxonomic breakdown of skills did not convey much information about dependencies among skills across checklist areas. Also, in the case of oral discourse, the checklist could not capture dynamic dependencies across several conversational turns of discourse. Therefore we must qualify that the skills checklist is not intended as a conceptualization of communicative competence but rather as an operationally oriented list of discrete

skills entering into communicative competence and the communicative process. As linguists from various theoretical perspectives have noted about language, communication, i.e., its integrative nature, goes beyond the sum of any identifiable discrete parts. Savignon (1983), Garvin (1978), Gumperz (1982), and others have noted that the gestalt nature of language, language usage, and communication is realized as a dynamic process, one that a taxonomy consisting of hierarchically arranged discrete, nonoverlapping categories cannot capture. In any case, despite the obvious inadequacies in a checklist approach to evaluating communicative competence, there are advantages in compiling such a list for research on proficiency testing. In this project, such a checklist is justified since it can serve as a useful analytical tool in looking at discrete skills that are tested intentionally and/or incidentally by a discrete point test such as TOEFL.

The identifiable discrete skills entering into natural communication in all of the various communicative and social contexts of language use are perhaps infinite. Consequently, those suggested in our skills checklist (Appendix D) comprise a select list of those skills that seem most salient and significant for a nonnative English speaker to learn in order to be adequately equipped to communicate in English in an American academic context. For the most part the skills selected were based on different types of meaning criteria --referential meaning, conceptual meaning, structural meaning, interactional meaning, attitudinal meaning, propositional meaning, etc. However, in conformity with the nature of English and its uniqueness compared to other languages, at some points the selectivity was based on patterns of organization or structure that may only be related indirectly to meaning. The selection of skills to be included in the checklist was based primarily on the uniqueness of English in American academic contexts. The designers of the final checklist drew heavily on their knowledge of other languages and cultures in the world, years of experience in teaching English to nonnative speakers, and sociolinguistic research. Above all, this skills checklist should only be construed as a suggested operational taxonomy of some discrete and salient skills foreign students must acquire and not as a definitive, comprehensive list of all that it takes to be communicatively competent.

3. Description of the Competence Skills Checklist

This section presents a general description of the major skills represented in the checklist (Appendix C) along with the rationale for the categorization suggested at each level in the taxonomy. The relevant portions of the checklist are cited in parentheses. Capital letters A, B, or C in parentheses indicate which of the general competencies is alluded to; digits following capital letters identify particular competencies or classes of competencies.

3.1 Grammatical Competence (Section A). This component refers to mastery of the linguistic code of English for which numerous grammars have been written, e.g., Francis (1958) Structure of American English, or, more recently, Leech and Svartvik (1975), A Communicative Grammar of English, and many others. Our skills checklist focuses on five basic dimensions of grammatical competence

that are discussed in depth below: pronunciation, script or written symbols, lexicon or vocabulary, word formation or morphology, and sentence formation or syntax.

Pronunciation (Section A.1)

Included in this section are skills related to the articulation of words as they combine in connected speech as well as the melodic accompaniment in prosodic patterns signaled by rate, stress, intonation, range of pitch, height of pitch, and pauses (Bolinger & Sears, 1981). Regarding the articulation of words or lexical items, a categorical distinction is drawn between those cases in which words maintain their full pronunciation form, for the most part, in connected speech (Sect. A.1.1) and those that are presented in a modified form in speech (Sect. A.1.2), reflected in such natural processes in English as vowel reduction, consonant cluster reduction, vowel deletion, palatalization, and contraction. There are pragmatic implications for making the distinction between Section A.1.1 and A.1.2 since rate of speech, degree of informality or casualness, and social context often trigger those linguistic processes mentioned in Section A.1.2.

While prosodic patterns in English are extremely complex, they have been known to play a significant role in learning how to communicate in English, given a non-English background (see Gumperz & Kaltman, 1980). There is much variation and even disagreement among researchers as to the patterns and functions of prosody in English; we have emphasized three overall meanings that prosodic phenomena participate in signaling in English: (1) neutral reference, with no extra emotional or emphatic message implied (Sect. A.1.3); (2) contrastive or emphatic meaning, where the opposite of the neutral reference is implied, marked by a variant in stress or pitch within an overall pattern (Sect. A.1.4); and (3) emotive or attitudinal meaning, marked by tempo, range of pitch, height of pitch, pauses, and other prosodic patterns (Sect. A.1.5). It is perhaps obvious at this point that the skills we have selected extend beyond an interest with denotative meaning to also include connotative meaning and the full range of meaning entailed in natural communication--even that at a more personal, emotional level. This orientation is in keeping with sociolinguistic research that describes communication as the negotiation of meaning based on inferences concerning the social and affective context of communication (Gumperz, 1982). Our treatment of prosodic patterns is not designed to describe in detail prosodic functions induced by the occurrence of particular phonological and syntactic structures in speech (Avery, Ehrlich, & Yorio, 1983). Such a description of prosodic functions would be valuable as an augmentation and refinement of the current work.

Script (written symbols) (Section A.2)

Skills selected within this section include graphic symbols and their combination as they are used to convey meaning in a written text. To be more comprehensive, we have selected both those symbols that are arranged into spelling patterns to form words (Sect. A.2.1) and those that signal various

types of grammatical, structural, and lexical meaning with a text (Sect. A.2.2 and A.2.3). Of those graphic symbols conveyed by the letters of English, there are those that reflect the spelling patterns found in isolated words (Sect. A.2.1.1) and those that reflect modifications caused by words entering into a sentence sequence (Sect. A.2.1.2), e.g., the problematic patterns of "plan" vs. "planned" or "cry" vs. "cried."

Other graphic symbols (e.g., punctuation symbols, slashes, underlining, indentation, capitalization, italics, boldface type) serve to mark a multitude of different meanings in written English text. Teachers are most familiar with the use of these symbols, particularly punctuation marks, to mark structural meaning or structural grouping in written texts (Sect. A.2.2); they are meant to aid the reader in selecting the salient pieces of information embedded linearly in the text. There is less analytical description about other uses of the array of graphic symbols to convey different types of meaning. We have suggested the following categories of uses of some of these graphic symbols: (a) representing modifications of full word forms (Sect. A.2.3) as in the case of contractions, for example; (b) marking specific concepts (Sect. A.2.4); and (c) marking emphatic or contrastive meaning within textual units (Sect. A.2.5). We have selected these particular uses of graphic symbols as a potentially comprehensive list of various types of meaning conveyed in written texts with the help of symbols other than letters. For example, the use of quotation marks, which serve to mark a variety of different meanings ranging from direct speech to conceptual meanings such as identity (as in nicknames--Bill "Skip" Smith or titles--in "Passages...").

Lexicon (vocabulary) (Section A.3)

Included within this skill category are the use of content words used with their neutral or literal meaning in context (Sect. A.3.1) as well as those content words used in a way that extends beyond their given literal meaning. At least two categories are distinguished among nonliteral uses of content words: (a) idiomatic expressions (Sect. A.3.2) or those instances in which words are frozen together to function as a semantic whole whose meaning cannot be determined by combining the meanings of the parts; and (b) metaphorical uses of words or expressions (Sect. A.3.3). Some very prominent types of idiomatic expressions in English that can be problematic to a learner of English include compound nouns (Sect. A.3.2.1) and compound (two-three word) verbs (Sect. A.3.2.2) as well as numerous others that have resulted for various historical reasons (Sect. A.3.2.3). Because of the complexities involved, our analyses of idioms were not designed to examine syntactic and other structural constraints that are implicated in the appropriate use of idioms (Yorio, 1980). Category A.3.3 (metaphorical uses of words) refers to unique uses of literal words to extend their meaning and must have a clearly identifiable literal and nonliteral reference since their interpretation depends upon the association of the two references, as in: "She is the flower of his life." Those metaphorical uses that were no longer recognized as unique or nonconventional were categorized under idiomatic expressions, e.g., "in one ear and out the other." Content words make up only that part of the

lexicon of which English speakers are most conscious and aware. Function words in English serve to bind content words together into complex patterns of meaning, often termed "grammatical meaning" by ethno-semanticists who have developed frameworks for uncovering this meaning (Mathiot, 1970). Typically, function words serve to mark specific concepts and relationships, as has been detailed by Munby (1978) and by Leech and Svartvik (1975). The most typical concepts marked by function words are included in the skills checklist (Sect. A.3.4.1-A.3.4.10). These include location and direction; definiteness and indefiniteness; quantity and amount; time, frequency of occurrence and span of time; means or instrument; possession; comparison and degree; and negation.

Morphology (word formation) (Section A.4)

In addition to function words, prefixes and suffixes attached to words serve to signal grammatical meaning, as is the case with inflection (Sect. A.4.1), or to mark the class membership of the content word involved, which may or may not be the case with derivation. We have included all types of derivation, provided they occurred in a context showing the relationships shared between two words having the same stem or root (Sect. A.4.2).

Section A.4.1 includes the major conceptual meanings that inflectional suffixes in English participate in marking either completely or partially: number, as in plural nouns (Sect. A.4.1.1); possession, as in possessive nouns (Sect. A.4.1.2); person, as in third person present verbs (Sect. A.4.1.3); various simple and compound tenses (Sect. A.4.1.4); and comparison and degree, as in comparative and superlative constructions (Sect. A.4.1.5).

Sentence formation (Section A.5)

Linguistic research in the past two decades has made us aware of the complexity of patterning in English at this level. It would be an impossible task to list all of these in this skills checklist; therefore, we have opted to select salient patterns that might prove problematic to a nonnative speaker of English whose task is to obtain meaning from given sentences. We have begun with an assumption posed by some current theoretical linguistic views backed by psycholinguistic research (Clark and Clark, 1977; Fodor, Bever, & Garrett, 1974). According to this assumption, derivation of the meaning of sentences is based on the ability to recognize the logical subject(s), verb(s), and object(s) encoded in sentence constituents, as well as ability to recognize semantic relationships linking together the information found in sentence constituents. Given the embedding and recursive nature of English syntax, these are not easy tasks. Embedding and recursion present special problems for English language learners. For example, learners must acquire facility in detecting the syntactical and lexical markers for these phenomena and, in concert with these skills, they must develop appropriate strategies for testing alternative interpretations of the emerging left-to-right syntactic structure of sentences. Complex sentences display many clauses with numerous clausal relations, which can make the task of selecting the logical relations a confusing one for a nonnative speaker of English. Therefore, we have selected those grammatical structures that are most typical of English

along with those that might seem most problematic in the recognition of the logical constituents making up sentences.

In the first group of structures are some commonly referred to by teachers as "simple sentence types" (Sect. 5.1). These are structures in which words or phrases--not clauses--participate in forming the logical relations among constituents. We have selected those well-known patterns that exhibit variation in the ordering of their logical constituents, but a variation that is systematic and suggests some sort of grammatical meaning, e.g., a question, a command, a statement. Included within these are declarative, active statements (Sect. 5.1.1); questions of any type (Sect. 5.1.2); imperatives (Sect. 5.1.3); passive constructions with or without a stated agent (Sect. 5.1.4); and THERE constructions (Sect. 5.1.5).

Section A.5.2 includes those sentences that teachers typically categorize as compound sentences. They are sentences with two main clauses and no subordinate clauses and as such have actually two sets of logical constituents bound together by some conjunction or other linking device. On the other hand, Section A.5.3 includes sentences with a main clause, from which the logical constituents are drawn, and a subordinate clause that stands outside the main clause--either preceding or following it--and adds additional semantic information to that contained within the main clause, as in: "When he's ready, he'll call" or "He'll call when he's ready," or, in reduced form: "Ready, he called."

Considered as a separate category from the type of complex sentence described in Section A.5.3 are those sentences in which clauses are subordinated within constituent members of the entire sentence--often referred to as embedding in current linguistic theory. We have chosen to include several processes or broad classes within this section that have been described in detail in recent years. Again, for the sake of minimizing excessive detail in the checklist, we will overlook some distinctions. Included in embedded structures are noun phrase (NP) complementation or nominalization (Sect. A.5.4.1), where noun phrases function as unitary constituents, and relative clause formation (Sect. 5.4.2), where subordinate clauses participate as modifiers of subjects or object constituents. In the latter case, consideration must also be given to those instances in which this particular subordinate clause is overtly signaled by relativizers, e.g., who, which, that, versus reduced instances in which no signal is present, as in: "The man I saw was nice."

A final sentence formation important to consider on pragmatic grounds, since some have suggested it represents a stylistic-based motivation for use, is extraposition and other forms of focus shifting (Sect. 5.5). For our purposes, we will place here all those instances in which the logical object precedes any order of the other logical constituents, as in topicalization, e.g., "That I've had enough of." or extraposition,

O S V

e.g., "What I think is that this is a difficult assignment."

O V S

3.2 Sociolinguistic Competence (Section B). We have relied heavily on the initial set of skills provided by Canale (Appendix A) in the development of this portion of the checklist. We have made rough attempts to expand and modify this set using various research orientations from work done on the ethnography of speaking and speech acts.

Section B.1 suggests some of those dimensions on which cultures define appropriateness when various ways of expressing the same information are available. We have selected some of the analytical factors that can be used to explain the way in which a message is presented, especially those that are socially defined. These factors include (a) status and power relationships between participants involved in communication (Sect. B.1.1); (b) topic (Sect. B.1.2); (c) setting and scene, whether formal or informal (Sect. B.1.3); (d) channel or mode of presentation, whether spoken or written (Sect. B.1.4); and (e) genre, or the general structure of presentation and format of the communication (Sect. B.1.5).

While Section B.1 includes skills that focus on the speech patterns and structuring of information from the perspective of the overall speech event or interactional event, Section B.2 includes those skills relevant at a smaller unit of analysis level, which is now termed "speech acts." It includes a selected list of possible ways in which utterances can be used within a given speech event or interaction, and the units are organized into categories of illocutionary acts or the attempt to accomplish some communicative purpose (Austin, 1962). Section B.2.1-B.2.20 lists a variety of illocutionary ways in which utterances or written pieces of language might be used communicatively. In order to capture crucial aspects of speech acts in natural communication, we have made a distinction between direct forms, in which the speech act matches the literal interpretation of the segment, and indirect forms, in which the speech act has a different meaning than the literal interpretation of the segment analyzed, e.g., where "It's hot in here," might serve as a directive to initiate the listener to open the window.

Section B.3 includes those uses of expressions that Malinowski (1944) has referred to as "phatic communication." In these instances, there is little meaning implied other than enacting a social ritual in a speech event or in marking the social role or social presence of others in a speech event. These are basically formulaic expressions with routine usage but minimal referential meaning.

Section B.4 lists some of the ways in which an indirect perspective toward a communicative event or an emotive or attitudinal orientation toward a discourse topic might be signaled either through grammatical forms or intonation. Some of the particular attitudes or emotive meaning implied that are included in Section B.4.1-6 concern sarcasm, ridicule, defeat, frustration, criticism, and doubt. Finally, Section B.5 suggests those linguistic forms or speech modes that are not typically found in broadcast or standard written English and that represent nonstandard dialect forms.

3.3 Discourse Competence (Section C). This portion of the skills checklist is concerned with the way in which communicative events--be they spoken or written--are bound together to form a whole. In keeping with current work on textual analysis, we have focused on those linguistic tools that can be utilized like thread in a cloth to bind the meaning of a text together as a whole. These have been labeled cohesion devices (Sect. C.1). Since a text is also held together by an ordering of ideas or propositions—coherence--we have included some skills that relate to the organizational structuring of given texts (Sect. C.2) in both spontaneous, unplanned conversational discourse (Sect. C.2.1) and in planned, preedited discourse--whether written or spoken (Sect. C.2.2). Our observations suggest that preplanning, as opposed to spontaneity with no preplanning, results in different structuring of the ideas in any given text, and thus would seem to have important implications for an academic use of English.

Section C.1 suggests some linguistic devices used to bind a text together according to the nature of the device itself: some are lexical (Sect. C.1.1) (e.g., pronoun reference, synonymy, word repetition, sequence markers, and so on), and some are syntactic (Sect. C.1.2) (e.g., ellipses, parallel structures, conjoined clauses). These devices have been organized in the skills checklist according to their formal nature because they can bind information in a variety of complex and different ways across different units within a text and in different manners across types of text. For example, pronoun reference can be used to bind elements of a sentence, paragraph, essay, or even book. The simplest way to handle this open variety of ways in which cohesive devices operate within the context of the skills checklist was to specify occurrence of these cohesive ties according to their immediate cohesive function at the sentence level or at the level of larger units of text. Sentence-level devices used to mark cohesion (Sect. C.1.2) were organized according to their functional nature in this process to mark conciseness (Sect. C.1.2.1); continuity (Sect. C.1.2.2); semantic relationships between clauses (Sect. C.1.2.3); and emphasis (Sect. C.1.2.4).

Coherence (Sect. C.2) represents those skills dealing with the organization of ideas within any given text beyond the boundary level of individual sentences. Most of those patterns in conversational discourse suggested by Munby (1978) have been utilized, including initiating the discourse, maintaining the discourse, and terminating the discourse (Sect. C.2.1.1-C.2.1.3). Since sociolinguistic research is still in the process of discovering the various mechanisms by which each of these skills is achieved in natural, conversational discourse, it was extremely difficult to operationalize them. There are no doubt many complex cues involved in the initiation, continuation, or termination of discourse drawn from the prosodic, kinesic, linguistic, and nonverbal systems of English that are important in American academic contexts and that vary from other English-speaking academic contexts throughout the world. As the literature has noted (e.g., Gumprez, 1982), these cues are quite often symbols of culture that operate at an extremely unconscious level in the communication process.

Kaplan (1966) suggested many years ago that planned, written discourse also exhibited cultural patterns in its organization of ideas. We

have suggested some organizational patterns that predominate written, expository texts in American academic contexts (Sect. C.2.2). Although there are many possible patterns of organization that could be utilized, as evidenced by numerous textbooks on written expository prose, we have used the framework developed and field-tested by Penfield for use with foreign students in Essay Writing for the African World (1979). Section C.2.2.1-C.2.2.8 suggests ways in which the predominating or salient organizational pattern of a given text might be categorized semantically, relying strongly on some typical rhetorical modes. Included are classification (Sect. C.2.2.1); illustration (Sect. C.2.2.2); definition (Sect. C.2.2.3); process or event-ordering (Sect. C.2.4); description (Sect. C.2.2.5); comparison (Sect. C.2.2.6); cause/effect (Sect. C.2.2.7); and factual development--where none of the previous categories predominate and facts are simply chained together (Sect. C.2.2.8).

3.4 Coding of TOEFL Items. Penfield and Duran separately coded every item on the TOEFL test found in Appendix A. The initial goal was to enact fully independent coding so as to allow a comparison of agreement in coding. This did not prove entirely feasible. Duran found it necessary to study Penfield's coding of a few items in each section prior to beginning his own independent coding. This strategy was necessary since initial attempts at establishing coding guidelines working with a separate set of TOEFL items not found in Appendix A had not sufficiently resolved questions on interpretation of the checklist. Duran's agreement with Penfield's coding was nonetheless quite high, over 80 percent, though Penfield's coding revealed a higher occurrence of some phenomena, e.g., grammatical phenomena related to pronunciation. It seemed clear that Penfield's more extensive knowledge of English structure, background in phonology, and experience in teaching English as a foreign language contributed to her sensitivity as a coder. A decision was made to use Penfield's coding as the criterion coding to be discussed in this report. One exception to this involved the coding of sociolinguistic information for items drawn from Section 1 of TOEFL. It was decided to use Duran's coding in this case because it proved to be more sensitive to plausible speech acts embodied in utterances.

The coding of TOEFL items against the skills checklist involved coding up to three separate pieces of information for each item. These pieces of information included (a) the language stimulus (or stem) upon which an item was based, (b) the question (if there was one) asked of examinees for an item, and (c) the correct response option. Incorrect response options were not coded.

Appendix D presents a summary of the coding that resulted. The summary identifies when a characteristic occurred once in the language stimulus, question, or correct response to a question; the fuller data upon which the summary is based identify exactly where characteristics were found within items. Appendix D aggregates this information according to each item type within the various TOEFL sections. The disaggregated data are not presented in this report because of bulk and the difficulty most readers would have in

interpreting it. (The full coded checklist is six times larger than Appendix D. While this more detailed coding is not contained in an appendix, results from it are discussed in the "Findings" section of this report.)

4. Procedures for Analyzing Performance Features

The sample TOEFL test in Appendix A was examined in light of the eight test performance factors cited by Bachman and Palmer (forthcoming). These factors, as discussed in Section II, include:

- o psychophysiological skills in test taking
- o representation of knowledge
- o language use situations
- o context and message
- o artificial restrictions
- o monitoring factors
- o affective factors
- o strategic factors

The initial plan of research was to formulate definitions of these factors in terms of a numerical ordinal scale that could be used to rate the importance of each feature for all the TOEFL items and, in summary form, for each TOEFL section and item type. This strategy did not prove feasible to the project staff person (Duran) who was responsible for establishing the rating procedures. It was judged that such an interpretation of TOEFL in light of Bachman and Palmer's eight factors could not be carried out definitively in a first attempt, but that there would be much value in presenting a preliminary discussion about the relevance of the eight factors to TOEFL sections and item types. It seemed that such a discussion was necessary to clarify the relevant notions of performance that needed attention. The resulting analysis is presented in the section of the report detailing findings of the study.

5. Procedures for Evaluating the Relevance of TOEFL Item Content to Academic and Social Language Use Contexts

The purpose of this evaluation was to capture the extent to which TOEFL items reflect representative language that examinees might encounter in criterion settings. Presumably, these criterion settings would encompass the social and academic universe encountered by foreign students upon entering a North American university. The goal of analysis was to develop ordinal ratings of the extent to which various TOEFL items manifest social and academic content. Penfield and Duran judged that the sociolinguistic competence section of the Checklist of Competence Areas for Describing TOEFL (Section B) already addressed the notion of content and that an alteration of the coding of the sample TOEFL against this portion of the checklist might accomplish the task at hand. Three ordinal rating dimensions for the category topic and authenticity were developed: academic living, academic content topic, and social naturalness. Academic living referred to the degree to which items were judged to reflect content and language use that students

might encounter in everyday college life in or out of the classroom. The academic content topic dimension referred to the degree to which items might be judged as having topics that could occur as part of formal instruction at the college level. The dimension of social naturalness referred to the degree to which items in the listening comprehension section of TOEFL were judged to reflect language that foreign students might hear in everyday social interaction. The three dimensions overlap in the sense that a given TOEFL item could be judged to convey information about all three simultaneously.

Establishing an ordinal scale and criteria by which to rate TOEFL items on these three dimensions proved quite problematical. This was so because individual TOEFL items in and of themselves tend to provide a limited sample of language and a limited notion of a speech situation. Nonetheless, after an initial abortive effort was made to establish a three-point scale for each of the three dimensions, a final scale was established. On this scale, a rating of "3" indicated that an item conveyed a high and clear level of authenticity with regard to the dimension in question. A rating of "2" indicated that an item could have been an authentic sample of language, but that this judgment required a rater to imagine a language use situation or topic that was not clearly marked by the actual language of an item. A rating of "1" indicated that an item did not manifest authenticity on a dimension or that the language making up an item did not provide enough information on which to base a judgment for a given rating dimension. Based on this procedure, ratings were computed by two independent raters along each dimension for each TOEFL item. Also, average scores based on each test section and item type within section were computed. Interrater indices of agreement were also computed and examined. The raters in question were research assistants to the project.

6. Procedures for Preliminary Description of the Proficiency Level Represented by TOEFL Items

In this part of the study, Charles Stansfield and Judith Liskin-Gasparro overviewed items on the sample TOEFL test using the Interagency Language Roundtable (ILR) scale. The scale was originally developed by the Foreign Service Institute (FSI) during the 1950s as a means of measuring the language proficiency of government employees, especially Foreign Service officers. Subsequently, the scale was adopted by other government agencies concerned with language training, including the Civil Service Commission, the Peace Corps, and the armed forces. It is also used by the armed forces of NATO member nations. (For more information about the scale and the oral proficiency interview associated with it, see Clark [1980].)

The ILR scale used in this study to describe TOEFL items is based on the above-mentioned FSI scale. At the time of this writing, it was the most current version of the scale available. The scale was developed by the Interagency Language Roundtable, a consortium of government linguists representing the various agencies involved in language training and language testing activities. (Recently, the American Council on the Teaching of

Foreign Languages (ACTFL) and Educational Testing Service have participated in meetings of the ILR by invitation.) The ILR scale is a description of proficiency in the four basic communicative skills: listening, speaking, reading, and writing.

There are eleven levels on the scale representing proficiency ranging from 0 to 5. A "plus" level is assigned to a person who fulfills most but not all of the requirements of the next highest level. The plus levels range from 0+ to 4+. There is no score higher than a level 5, which is defined as "proficiency equivalent to that of an educated native speaker." Levels on the scale are defined in paragraph-length descriptions. The November 1983 version is reprinted in full in this report as Appendix E.

The ILR descriptions for each level are based on the ability to handle various communication situations. These increase in difficulty along the scale. Two examples are the ability to understand directions on how to get from one place to another (level 1, listening), and the ability to read all materials in one's professional or technical field (level 4, reading). Performance ratings on the scale reflect the ability to send or receive information in a wide variety of situations that can be associated with different levels of overall proficiency in a second language. A rating is assigned based on three criteria: the level of the text, the level of the task being asked of the examinee, and the level of the examinee's performance. Thus, the scale assumes that the ability to handle various situations effectively in a second language is hierarchical. The scale incorporates both linguistic complexity and situational competencies in its definition of communicative competence levels, and it is used by government agencies to assign a proficiency level to a communicative text. Because the scale is already being used to assign a difficulty level to a test in a test situation, it was decided to ask two persons familiar with the scale to apply it to TOEFL.

As noted earlier, in this study, Charles Stansfield and Judith Liskin-Gasparro applied the ILR scale to TOEFL. Both are familiar with the scale and certified to train other examiners in its use. They jointly evaluated items on the sample TOEFL form to gain a better understanding of the proficiency level represented by the items. They individually reviewed the form and then jointly discussed it. All items in Sections 1 and 3, and nearly all items in Section 2, were discussed and assigned ratings on the scale. Four and one-half hours of discussion were tape recorded. Subsequently, the examiners prepared the observations given in the "Findings" section regarding the scale and the TOEFL test.

IV. FINDINGS

1. Plan of Discussion

The characteristics of the sample TOEFL test in Appendix A will be discussed in terms of each of the four framework components for analyzing the validity of a language proficiency test. Findings with regard to each framework component will be discussed in the same order as used in Sections II and III of this report. At times the reader may notice that there is an overlap in the discussion of findings across framework components. This confluence is reflective of a natural overlap in issues. The various framework components were not intended to be mutually exclusive, but rather to be complementary in their scope and analyses. The task of describing the TOEFL test in light of the framework is not strictly an act of reporting objective data. Considered judgments are required that stem from the investigators' experiences and views on the structures and phenomena being described. This need to interject judgment is a characteristic of many forms of linguistic analysis and of test content validity research.

Figure 1 outlines the order of discussion of the findings and the subsection number involved. This outline is useful for understanding the logical organization of the discussion.

2. Description of TOEFL Based on the Checklist of Communicative Skills

2.1 Summary of Findings for the TOEFL Test as a Whole. The data under discussion here are entered on the skills checklist given in Appendix D along with instructions on how to interpret tabular entries. The terms used to describe various phenomena in this section are linked to the same terms occurring in the skills checklist. Please note that entries on the skills checklist are sensitive only to the simple presence or absence of skills on each item making up TOEFL test sections; these tallies are not sensitive to the total number of occurrences of skills within individual items.

The data in Appendix D reveal some broad patterns regarding the occurrence of communicative skills on various TOEFL test sections. In some instances we found a number of shared skills across all sections, but in other instances we found some patterns of skill occurrence that were unique to a given section and its item types. An important conclusion to state from the outset is that the range and complexity of skills required on various sections of the TOEFL test are directly related to the amount of language used and to the semantic and textual complexity of TOEFL items. The more language used, and the more authentic this language, the greater the number and kinds of communicative skills required of examinees. The reader should keep in mind that this conclusion should be tempered by the fact that TOEFL has an assessment purpose that is more limited in scope than assessment of the full range of communicative skills discussed here.

Figure 1

Outline of the Discussion of Findings

2. Description of TOEFL Based on the Checklist of Communicative Skills
   - 2.1 Summary of Findings for the TOEFL Test as a Whole
   - 2.2 Summary of Findings for Listening Comprehension, Section 1
     - 2.2.1 Section 1: Statements, Part A
     - 2.2.2 Section 1: Dialogues, Part B
     - 2.2.3 Section 1: Minitalks/Extended Conversation, Part C
   - 2.3 Summary of Findings for Structure and Written Expression, Section 2
   - 2.4 Summary of the Findings for Vocabulary and Reading Comprehension, Section 3
     - 2.4.1 Section 3: Vocabulary
     - 2.4.2 Section 3: Reading Comprehension
   - 2.5 Concluding Comment on the Checklist Findings

3. Findings Regarding Test Performance Factors Cited by Bachman and Palmer
   - 3.1 Overview of Issues
   - 3.2 Psychophysiological Skills
   - 3.3 Representation of Knowledge about Language and Monitoring Factors
   - 3.4 Language Use Situations
   - 3.5 Context and Message
   - 3.6 Artificial Restrictions
   - 3.7 Affective Factors
   - 3.8 Strategic Factors

4. Findings Regarding the Relevance of TOEFL Item Content for Academic and Social Language Use
   - 4.1 Rating Agreement
   - 4.2 Ratings of the Characteristics of Section 1 Items
   - 4.3 Ratings of the Characteristics of Section 2 Items
   - 4.4 Ratings of the Characteristics of Section 3 Items

5. ILR Scale Description of TOEFL Proficiency Level
   - 5.1 Section 1: Listening Comprehension
   - 5.2 Section 2: Structure and Written Expression
   - 5.3 Section 3: Reading Comprehension and Vocabulary

6. Conclusion

Some general tendencies for the entire TOEFL test to require skills found on the checklist can be noted when examining the overall frequency of occurrence of the more general classes of skills. For example, in Section 1 at the pronunciation level, recognition of use of tempo, range/height of pitch, or pauses to mark emotive or attitudinal meaning were just as frequently required of examinees as was recognition of stress and intonation to mark neutral, nonemotive reference. Each occurred in somewhat fewer than one-third of the total items in Section 1. In addition, it is equally interesting to note that examinees' need to recognize the modification of lexical items in connected speech in Section 1 occurred just about as frequently as examinees' need to recognize the modification of inherent spelling patterns of words in the remaining sections of the test.

Because the majority of the TOEFL test is presented in a written mode and all correct answer options are presented in writing, it is not surprising to find that recognition of graphic symbols plays an especially crucial role in TOEFL test performance. The table below, constructed from information found in Section A.2.2 of Appendix D, lists the percentages of items in the overall test requiring recognition of graphic symbols performing the functions or encodings listed.

To mark structural groupings:	
clauses	29%
phrases	26%
Speech-based modifications:	
(e.g., contractions)	21%
To mark concepts:	
mathematical	7%
formal identity	31%

Attention now turns to data on some of the important lexical and morphological features of the language occurring in TOEFL items. The data are tabulated in Section A.3 and A.4 of Appendix D. Compounding seems to occur fairly frequently, with 58 percent of the 150 TOEFL items requiring recognition of one or more noun compounds and 31 percent requiring recognition of verb compounds. Regarding inflectional markings occurring on nouns and verbs, recognition of linguistic encoding of number seemed to be a much-required skill, as indicated by the fact that nouns marked for plurality occurred in 73 percent of the items and marking of third person singular present tense verb endings occurred in 53 percent. By far, the most typical inflectional tense markings, other than third person, were regular and irregular forms of the simple past; this occurred at least once in 46 percent of the items. Present perfect marking occurred next most frequently (28 percent of the items); past perfect and present progressive rarely occurred (7 percent and 5 percent, respectively). Qualification, case relations, and other devices to relate content words together play an important role in English language communication, and this importance was well represented by TOEFL items. Marking of specificity—definite or indefinite and often both—occurred at least once in almost every item. Other frequencies of occurence of those functions in the 150 items across TOEFL sections in the test form analyzed are given below:

Location/direction	59%
Quantity/amount	51%
Possession	47%
Time (frequency/span)	39%
Negation	37%
Comparison/degree	20%
Instrument/means	5%

Turning to Section A.5 of Appendix D, sentence formation in the TOEFL test typically exhibited the declarative, active pattern of subject-verb-object (S-V-O); this pattern occurred at least once in 88 percent of the items. Two other patterns also typically occurred: passives, at least once in 45 percent of the items; and direct, complete questions, at least once in 30 percent of the items. This points out the special importance of examinees being able to construct or at least interpret ordering of sentence elements in order to do well on TOEFL. Two types of within-sentence subordination of clauses also occurred somewhat frequently: (a) complex sentences with one or more subordinate clauses outside the main clause--at least once in 35 percent of the items; and (b) relative clauses--at least once in 34 percent of the items.

The summary tallies involving recognition skills related to sociolinguistic competence required in working TOEFL items (Section B of Appendix D) bring some important observations and raise some specific questions. For some of the sociolinguistic skills, many TOEFL items were very difficult or impossible to code with consistency across coders. For example, recognition of status/power relationships, topic, and setting/scene based on the language of items was often difficult or impossible to code unambiguously. The three skills that are cited are situational and socially derived in real-world language contexts, but TOEFL items focusing on discrete linguistic skills or neutral, written forms of communication did not typically provide enough contextual information to allow an unqualified judgment about the elicitation of these skills in understanding the language of items. Examples that especially illustrate this coding difficulty are the isolated statements making up items occurring in the vocabulary part of Section 3; by design, none of these statements is connected to another or to any other discourse context presented. Consequently, coders interpreting items based on real- world criteria for language use had to infer possible situational contexts from linguistic stimuli in isolation. And often--not surprisingly--one particular statement could be judged to entail multiple situational contexts at varying levels of generality for any one native-speaking coder. In short, the discrete-oriented, shorter TOEFL item types did not provide the coders with enough textual or situational context information to enable them to unambiguously identify the setting, or topic of discourse, and the social roles and role relationships among interlocutors depicted. This finding should be tempered by the fact that items in some TOEFL sections are designed to be content independent of each other in order to reduce the chances of item content bias. The longer, more integrative item types provided lengthier and structurally more complicated texts but, most usually, were limited in their interpersonal and social uses of language relative to authentic everyday communication.

As will be seen in a subsequent section, however, the two independent judges who rated the authenticity of items found that the longer, more integrative items bore greater authenticity to actual language use.

Consistent with the findings to be presented in the section on test performance factors (Section 3), the Appendix D, Section B.1.5, data on genre indicate that, in our judgment, the isolated segments of language that make up many TOEFL test items, particularly shorter items, show a limited degree of commonality with natural communication. Overall, these data show, for example, that 13 percent of TOEFL items involved segments of conversation, while isolated statements (e.g., item language stimuli in isolation from a specified context) and limited functional statements (e.g., isolated answers to isolated questions) occurred in 71 percent and 43 percent of the items, respectively.

In authentic language use, especially that involving language at the interpersonal communication level, one typically finds a large variety of communicative purposes enacted. Most of the TOEFL test (because of its problem-solving format) emphasizes examinees' recognition of direct requests for information (39 percent of the items) and direct giving of information (97 percent of the items). A wide variety of other speech act functions typical of everyday interpersonal encounters are exhibited in Section 1, by inferring the social and cultural context indirectly signaled by the language of items. In contrast, Sections 2 and 3 (see Section B.2 of Appendix D) feature skills requiring recognition of direct giving and requesting of information.

Attention is now turned to a summary of discourse competencies required in understanding TOEFL items (see Section C of Appendix D). Although nonreferential speech act functions employed in authentic language use were inferred to occur frequently only in Section 1, other discourse recognition skills typical of those required to understand textual or narrative development were prevalent across all sections of the test. Ability to recognize lexical cohesion devices was a frequent requirement on TOEFL as a whole, which is illustrated by the percentage of items demonstrating the use of cohesive devices listed:

Pronoun reference	59%
Word repetition	40%
Conjoiners	39%
Synonymy	33%

Examinees' need to recognize sentence-level devices used for cohesion and continuity also were often found; need for such skills occurred in 49 percent and 31 percent of items, respectively. The following list presents the percentages of items requiring recognition of various semantic, interclausal relationships involving conjunctive adverbs.

Contrast	28%
Result/conclusion	21%
Time	20%
Addition	18%
Condition	14%

It almost appears that occurrence of the forms of semantic conjunction listed were equally distributed intentionally.

It was expected by the designers of the skills checklist that, since the written portion of TOEFL relies on texts drawn from encylopedic sources, textbooks, or academic journals, certain discourse structures frequently occurring in planned and edited texts would be found (e.g., classification, illustration, definition). All of these forms of text development are prevalent in student textbooks for native speakers and nonnative speakers, and they do exhibit some of the semantic rhetorical forms found elsewhere in English. Therefore, it was of some surprise to find that only two such patterns were exhibited in Sections 2 and 3 of the sample TOEFL test; these involved event sequence development and an expository text category termed "factual development." It is difficult to explain why no other forms of text development occurred, since the sample TOEFL had five reading passages in the reading comprehension portion of Section 3. It is likely that the sample TOEFL test analyzed is not representative of TOEFL tests in this regard.

The foregoing discussion was intended to give a summary overview of important communicative skills required and not required on the TOEFL test as a whole, as reflected by the skills checklist of Appendix D. The discussion was not intended to be exhaustive, but rather was guided by some important trends that emerged in the data. With this summary in mind, we now turn to a more detailed examination of the communicative skills required within each TOEFL section and item type. Discussion of each section will comment on important characteristics of the item types making up the section; the data discussed are based on a more detailed and voluminous breakdown of Appendix D data. This more detailed breakdown of data records the occurrence or lack of occurrence of every checklist characteristic for every item stimulus and correct response separately. Readers with limited interest in these details may elect to skip this next section and instead proceed to Section 3, concerning with other factors affecting performance on the TOEFL. Readers with a serious interest in the report, however, should read the next section.

2.2 Summary of Findings for Listening Comprehension, Section 1. Section 1 consists of three parts, which differ most obviously in their length. The items in Part A, statements, consist of sentence statements that occur in isolation of any particular context in which they might naturally be found. Part B, dialogues, consists of statements by two different speakers that form an interactional unit. Part C, minitalks/extended conversation, consists of multiple sentence statements that are bound together either in a textual unit to simulate a lecture or announcement, for example, or in an interactional unit, to simulate a multiple-turn conversation between two speakers. However, the linguistic nature of Part C, even when it appears to be a conversation, makes it appear in some ways more similar to written expository prose than to informal dialogue or interactional units such as those in part B. For example, the topic dealt with, the occurrence of complex sentence structures, and the reliance on historical and descriptive flow of information render any passages in Part C more typical of planned, edited discourse than of spontaneous, unplanned discourse so typical of conversations, dialogues, and even

many lectures. Impressionistically, one has the pervading feeling, especially *when listening to Part A and Part C*, that edited, planned prose typical of written English contexts simulating conversation is being read from a script and spoken. This quality of speech, notable because it is crisply enunciated with no unexpected pauses and with minimal ellipsis, is no doubt intentional on the part of the developers of the TOEFL examination. Such speech helps ensure that examinees have a maximum chance to recognize the meanings of utterances; it does not require examinees to utilize speech recognition skills that deviate far from the lexical and grammatical structure of the stimulus materials. Interestingly, a similar style of speech, alluded to as "foreigner talk," has been found to occur when native speakers make a special effort to communicate with nonnative speakers (Hatch, 1983). It should be recognized, however, that authentic everyday speech, outside of situations such as those involving "foreigner talk," formal academic lectures, etc., seldom manifests the crispness of pronunciation, absence of ellipses, absence of false starts, and uniform grammatical correctness of speech found on the TOEFL test.

Because strategic competence was not included in the skills checklist framework used to analyze TOEFL, attention was not given to some important aspects of natural communication involving this kind of competence. It was not required to any noticeable extent in understanding Section 1 material. Aspects typical of spontaneous communication, such as hesitation, repair, unfinished statements, paraphrasing or rephrasing in midsentence, interjection of side comments, and many others are part of everyday natural, spontaneous communication. The checklist does not include any of these skills that are indicative of spontaneous, communicative behavior. There is no loss in this omission since, in looking at the extended conversations and minitalks, none of these particular strategies/competencies occurs. However, to a native speaker, the samples of language occurring in Section 1 may not be considered entirely representative of spontaneous speech.

In addition, in the process of applying the skills checklist to Part B, another set of skills was not included that, in this case, would have been helpful analytically. In accordance with the current popular view of communication as intimately involving negotiation of shared meaning, whereby speakers participate in shaping and directing the structuring of spoken text, more distinctions than simply how to initiate, maintain, or terminate a discourse would have been helpful. More attention should be given to knowledge of the flow of communication within speech events (Duran, 1984; Yorio, 1980). In particular, we now would recommend distinctions/recognition skills pertinent to recognizing the flow of information in a dialogue, as occurs, for example, in the sequence of the following speech acts: a direct request for information, a direct response to such a request, clarification of previous information as a response, elaboration of information, and modification of or disagreement with previous comments, suggestions, or merely emphatic comments that have little referential value but much emotional value. Other descriptors of contingent discourse behavior might also have been useful. These skills

seem to play an important part in spontaneous interactional contexts whether they be brief, e.g. Part B, or extended, e.g., Part C. Although such competencies were not originally included in the skills checklist, some mention will be made of them in the discussion of Parts A, B, and C of Section 1.

Section 1 is the only TOEFL section requiring skills involving recognition of pronunciation. To some extent the pronunciation offered attempts to simulate real-life speech, as reflected by numerous uses of contractions; these occur at least once in 36 percent of the items. However, other typical modifications of isolated word-form pronunciation in connected speech--e.g., elision, vowel reduction or deletion, consonant reduction, and palatalization--occur much less frequently than they would be encountered in informal natural, English speech. The checklist data support the observation that the speech for all three parts of Section 1 is clearly enunciated. Because of the desired trait of maximal clarity of speech in a TOEFL testing context, it is lacking in assimilation patterns of spontaneous, everyday communication. When one analyzes other aspects of the linguistic content of the verbal statements in Section 1 of the test, the "careful" style, as Joos (1967) would refer to it, is presented phonetically but mixed with informal style, marked by use of idioms and colloquial vocabulary, occasional elliptical responses, contractions, and sentence formation pragmatically tailored for a particularly interactional event. For example, consider the stimulus portion of item 32 in Part B of Section 1:

A: "Didn't you tell Tom about the meeting?"

B: "Whatever I say to him goes in one ear and out the other."

In this example, there is minimal palatalization in "didn't you," minimal vowel reduction in either statement, and the contraction and idiomatic expression reflect informal style.

One final, phonetic-based comment can be made about many items presented in all parts of Section 1. Contrastive or emphatic word stress occurs at least once in about 50 percent of the stimuli for items. In Part A and B a very frequent occurrence of higher pitch patterns (level 3 or 4) occurs throughout the stimulus statements. The combination of stress and high pitch patterns on more than one segment of a statement renders an initial impression of overly dramatized speech. In the test form studied this was particularly noticeable for the female speaker (e.g., Part A, item 7), with three instances of third-level or above pitch. Perhaps an exaggerated speech style is a natural accompaniment to a careful speech style, where each word is precisely enunciated. However, the correct options for items usually involve neutral, nonemphatic, and noncontrastive referential use of language, while the actual stimulus statements they paraphrase often have attitudinal, emotive meanings that extend beyond mere referential, linguistic meaning as prompted by stress and intonation patterns.

2.2.1 Section 1: Statements, Part A. Following is a more detailed discussion of each item type making up Section 1, beginning with statement items. Idiomatic expressions in statement items are prevalent in stimuli and in correct response options; noun and verb compounds make up the most typical type. (About 50 percent of the item stimuli and 25 percent of the correct response options have one or more idiomatic expressions.) As far as function words are concerned, those used to mark either generic or specific concepts are present in more than 50 percent of the stimuli and 50 percent of the correct response options. Possession and time span concepts marked through function words are the next most frequent in occurrence (35 percent of both the stimuli and the correct response options). The most frequent inflection is the marking of simple past forms, regular and irregular; this occurs at least once in more than 50 percent of the stimuli and correct response options.

The most frequent type of sentence formation occurring is the simple sentence word ordering of declarative, active statements, prevalent in about 95 percent of the stimuli and correct response options. Other types of simple sentence word ordering, exemplified by structures appropriate for statement of questions, imperatives, passives, and existential THERE sentences, are rare, as are complex sentence patterns with subordination. Compound sentence word ordering also seems somewhat uncommon.

Because all of the stimuli and correct response options in Part A occur in unspecified social contexts, it is impossible to judge precisely which sociolinguistic recognition skills are required of examinees. There simply is not sufficient information about an interactional context to code isolated sentence statements with a high degree of certainty as to various sociolinguistic recognition skills. Even such skills as those related to topic recognition become difficult to assign on the basis of one isolated statement. Impressionistically, it appears that most statements are not marked as unquestionably pertaining to specific academic living contexts or academic classroom content topics. However, more detailed findings regarding topic are discussed in a later section of this report.

Speech act functions are difficult to code because of the lack of explicit information in statement stimuli regarding the social and interpersonal context of speech. Nonetheless, some judgments are possible when one relies on native social, and cultural knowledge of American life and, to some extent on knowledge of generalized international English usage. A little more than half of the stimuli for statement items give information simply and directly (55 percent); the remaining stimuli may indirectly advise (e.g., the stimulus for item 6), invite (the stimulus for item 8), seek approval (the stimulus for item 7), make a suggestion (the stimulus for item 15), etc.

Linguistic devices most common at the discourse level include pronominal reference (more than 75 percent of the stimuli and correct response options) and sentence level markers of conciseness, e.g., ellipsis, clausal reduction (about 10 percent of the stimuli and correct response options).

2.2.2 Section 1: Dialogues, Part B. Idiomatic expressions are even more prevalent in Part B, occurring at least once in almost two-thirds of the stimuli for dialogue items. The majority of these idiomatic expressions are noun and verb compounds. Since these dialogues attempt to simulate informal conversation, one would expect to find such a prevalence of compounding--a process typical of colloquial, informal, or slang language usage. The occurrence of idiomatic lexical items may account in part for the difficulty of dialogue items since they reflect a colloquial, more informal use of language to which many foreign students are not exposed in formal, technical-language training.

In contrast to Part A, in Part B simple past tense inflections rarely occur. A quick glance over all the stimulus dialogues and succeeding questions suggests an extremely frequent pattern not captured in the coding of the skills checklist, but nevertheless important to recognize as a characteristic of items. First, a majority of the dialogue stimuli consists of first- and second-person interchanges--singular or plural--while all of the questions are formed in third person present singular forms, using either "does" or "is" forms. Secondly, if one searches for a critical linguistic clue or for critical semantic content for answering the question posed about a dialogue, in almost every instance the critical clue appears to exist in the second speaker's statement (Penfield judged this to occur for 86 percent of the items). Therefore, if examinees focus primarily on the second speaker's statement, they have a better chance of selecting the correct option. It is this second observation about Part B that raises the question of whether this part is in fact testing communicative skills typical of a dialogue or whether it is simply testing comprehension of a target statement in a slightly more extended context than that given in Part A. It appears possible that the language skill focus of these items may be based more on accurate recognition of the second statement as a target rather than on recognition of an integrated interactional event as a whole.

Examinees' receptive knowledge of sociolinguistic norms is difficult to code, and this seems especially evident with regard to factors related to norms of appropriateness of usage and to the social relationships of interlocutors. There is limited information about the context of communication surrounding the occurrence of dialogues. Because of this and because of the limited information about context found in the conversational turns of a dialogue, the topics treated in these items generally do not seem to clearly, directly, and unambiguously represent academic living contexts or classroom learning content topics. The stimulus dialogue for item 34 is an exception. Perhaps the most interesting and unique aspect of Part B is the use of language to accomplish communicative purposes. Most stimuli for items in Part B are not simply direct requests for or offering of information. Rather, there are several other direct and indirect communicative purposes reflected in Part B language stimuli. Indirectly expressed purposes are no doubt more difficult for non-English speakers to recognize since they involve reliance on knowledge of sociocultural norms and contexts for their accurate interpretation. Some of the indirectly expressed messages in dialogue stimuli include

request for help (items 26 and 31); promise (item 26); complaint (items 25, 27, and 32); slight insult (item 32); suggestion (items 29 and 31); impatience or annoyance (items 30 and 32); advice (item 31); and denial (item 35). Perhaps the most difficult to code is the stimulus for item 35, in which both attitude and communicative purpose are expressed in a very indirect way that is also quite idiomatic and dependent on recognition of the tone of utterances.

Discourse competence recognition skills, that might be thought to be the most revealing, extend beyond the skills checklist to include concern for the turn-taking structure of dialogues, i.e., perception of the rights and obligations of interlocuters and their fulfillment of these rights and obligations. In dialogue stimuli about 20 percent of the statements made by the second speaker reflect a direct or an indirect response to the request of the first speaker; another 20 percent reflect a modification of the first speaker's comment. Other dialogues reflect elaboration of information, clarification, and suggestions offered by the second speaker in reaction to the first speaker's statement/comment.

2.2.3 Section 1: Minitalks/Extended Conversations, Part C. Part C displays a formal style of speech that closely approximates written expository prose in its vocabulary, sentence formation, and discourse structure; this formal character is called for by the TOEFL test specifications. Idiomatic forms are much less prevalent on this portion of Section 1, and this is especially noted in the reduced occurrence of noun and verb compounds. Many more competencies are required in understanding stimulus passages for Part C since the length and complexity of item stimuli are greatly increased over those in Parts A and B. The longer passages in Part C entail recognition of at least one occurrence or more of idiomatic and literal meanings of content words; literal meanings of various concepts marked by function words; and inflections for number, possession, person, and most tenses. Sentence formation recognition skills required of examinees in Part C include simple sentence word order and compound and complex sentence word order with embedded subordination, neither of which is required as much in Part A or B. The topics occurring in the stimuli passages of Part C are more identifiable as relevant to academic life or classroom content than is the case with the topics in Parts A and B. The longer passages often involve descriptive information about American geographic locations and history drawn, usually verbatim, from authentic texts. Because of their expository prose nature, most of the communicative purposes served by the connected statements comprising the stimulus passages of Part C items are simply direct communication of information. This is so even for the extended conversation in the last stimulus passage of Part C. Although this stimulus is comprised of ten conversational turns between two speakers, it consists almost entirely of information giving and requesting and involves only a limited negotiation of topic. Thus, it does not strongly resemble a spontaneous exploration between two speakers of a topic of mutual interest.

The relationship between test questions about Part C stimuli and correct response options is elliptical for about 50 percent of the questions. The occurrence of ellipses exercises examinees' skill in recognizing ellipses,

though the communicative context involved is identifiable as a formal multiple-choice testing context rather than as everyday communication. Some elliptical forms in questions and correct responses consist of one word; others, one phrase. In all cases, one part of the message lies in the question and the other in the correct response option. Whether correct responses are elliptical or full sentences, each begins with a capital letter and ends with a period to mark completion. This is also true of items in both Parts B and C of Section 1.

2.3 Summary of Findings for Structure and Written Expression, Section 2

Both item types in Section 2--structure and written expression--obviously require use of orthographic recognition skills, skills that are much more numerous and varied than those required in Section 1. For example, recognition of capitalization to mark specific concepts, such as formal identity, frequently accompanying proper names of people, places, nationalities, etc, appears at least once in almost 50 percent of the item stimuli of structure and written expression items combined. Graphic symbols to mark clausal groupings occur much more frequently in structure items (for about 33% of the items) than in written expression items (for about 8% of the items).

It is obvious that the most essential competencies or skills necesary for both parts of Section 2 are those concerned with recognition of vocabulary, word formation, and sentence formation. Correct response options seem based on conventional word orderings in English for a variety of contexts and the skills tested, including correctly recognizing an appropriate word in a sentence context; the units of a phrase required within a sentence context; the order of words or phrases required within a clause in a sentence context; the required order of words and identity of words making up an appropriate idiomatic expression within a sentence context; and the required presence or absence of inflection on words within a sentence context. The items in Section 2 require no attention to meaning developed across sentences.

In addition to the specific structural recognition skills that seem to be tested in Section 2, one can observe the prevalence of other aspects of grammatical competence that may not be tested but that exist nevertheless as part of the sentence contexts making up Section 1 items. For example, idiomatic expressions appear very often--principally compound nouns (42 percent of the time in both parts). Function words and inflections on content words are used in both parts to represent a variety of grammar-based concepts. For example, location/direction occurs at least once for 35 percent of the items in Section 2; marking of specificity occurs for almost all of the items; marking of quantity for 37 percent; marking of time for about 10 percent; possession for 37 percent; and comparison/degree for 12 percent. Inflections occur in almost all items, with plurality marked in 72 percent. Third person singular present tense and simple past tense were marked most frequently; each occurs at least once in 40 percent of the items.

With regard to sentence formation recognition skills, by far the most frequently required recognition pattern for items involves the S-V-O order of declarative, active statements; this pattern occurs for 77 percent of the items. Because of the format of Section 2 items, sentences following a grammatical question ordering of constituents never occur.

Sociolinguistic recognition skills were difficult to code for Section 2 because only isolated sentences were presented. Impressionistically, it seems that over half of the items in Section 2 deal with topics that are not linked to clearly identified academic living contexts or classroom content subject matter. However, one might judge that many topics refer to classroom content typical of secondary school or college courses in America, since so much reference is made to everyday American history or geography concepts. The findings about orthographic recognition skills required of examinees is related to examinees' knowledge of concepts that appear to occur frequently in items. Famous people, places, dates, historical events, books, or magazines that are typical of U.S. culture and background and taught in American public schools at the preuniversity and university levels appear frequently in the items in Section 2--at least once in an estimated 35 percent of the items. Noticeably, a number of references to persons, events, and entities in items utilize compound nouns that are typical of U.S. history subject matter or to American names for plants or animals.

Discourse recognition competencies required in Section 2 are minimal compared to recognition of sentence structure and vocabulary. Some recognition skills at the discourse level are required though, within the contexts of single sentences. These skills include recognition of lexical cohesion and sentence-level cohesion devices, e.g., use of pronouns, clausal reduction, parallel structuring, and transitional conjoiners binding semantic relationships between clauses.

2.4 Summary of the Findings for Vocabulary and Reading Comprehension, Section 3

Section 3 appears more meaning oriented than Section 2. The first part of Section 3, vocabulary, tests recognition of appropriate lexical meaning as manifested by words, idioms, or phrases; the second part, reading comprehension, tests for recognition of textually manifested meaning. The two parts differ in the size or length of the item stimuli, which explains why a larger number and variety of the checklist skills were found in the reading comprehension subsection than in the vocabulary subsection.

2.4.1 Section 3: Vocabulary. As in all parts of Section 2, recognition of written symbols or script recognition skills are necessary in decoding the stimulus and correct response options for vocabulary items. Note that there are no question stimuli for these vocabulary items. Some of the graphic symbols that commonly occur in vocabulary items include those used to mark concepts requiring formal identity. Capitalization and/or quotes to mark names of people, cities, states, rivers, dates, and so on occurs at least once in 37 percent of the vocabulary items. The attention paid to mark formal

identity in vocabulary items correlates highly with the nature of the topics dealt with in these items. About 60 percent seem to deal with topics that might be judged to be related to American history and geography. Most of the capitalization serves to mark reference to historical/geographical concepts that are part and parcel of American culture and history; this usage is intentionally built into these items. Just as we mentioned with regard to Section 2, these concepts do not appear in the correct response options, and they certainly do not themselves reference the discrete skills tested. Thus, it remains to be determined whether knowledge of these concepts helps examinees by contributing information that is useful in inferring the meaning of the targeted vocabulary items.

Other graphic symbols occurring in vocabulary items that somehow contrast in frequency of occurrence with those of the reading comprehension portion of Section 3 are those used to mark structural groupings. In vocabulary items, commas are used to set off phrases in 23 percent of the items, and to set off clauses in 10 percent of the items. In the reading comprehension portion, however, both uses of commas occur in all passages. For example, use of commas to separate clauses occurs in all reading passages used in the reading comprehension portion of Section 3. Such differences between the vocabulary and reading comprehension parts are no doubt due to the difference in length of the language stimuli presented, though the stimuli items in each section are drawn from authentic texts. In the vocabulary part only one complete sentence is presented per item, whereas in the reading comprehension many connected sentences are presented, and questions posed for the items present yet another sentence context.

In the vocabulary part of Section 3, the stimulus portion of items frequently manifests idiomatic vocabulary, especially noun compounds, which occur in almost 50 percent of the items. Function words are fairly prevalent in the items, with location/direction most often marked (in about 50 percent of the items). Thirty-three percent of the items include marking for quantity/amount; 13 percent for time/frequency of occurrence and span; and 10 percent for comparison/degree. Certain inflections are extremely prevalent in vocabulary items. For example, marking of noun plurality occurs in 27 out of 30 items and simple past verb inflection marking in 14 out of 30. By far the most typical ordering of logical constituents in the vocabulary sentence items is that of declarative, active statements: subject-verb-object occurs in 22 out of 30 items. Passive constructions seem to occur for vocabulary items as frequently as for Section 2, and more frequently than for any other portion of the exam. Such structures can be found in almost 40 percent of the items in both the vocabulary portion of Section 3 and the structure portion of Section 2. The fact that the length of the sentence stimulus is minimal or that the sentences used are not part of clearly established textual contexts may explain why embedding occurs infrequently and why complex sentences are few in the vocabulary portion; the vocabulary part does not have any compound sentences and an average of only 10 percent of each complex sentence type. This, however, is not the case with the reading comprehension portion of Section 3, as we shall mention later. The short length of vocabulary stimuli

does have an advantage: it helps ensure that examinees will not infer the meaning of the missing words solely on the basis of the meaning of words in the stimuli.

A few interrelated discourse skills are worth mentioning in relation to vocabulary items. Sentence-level cohesion devices are noticeable; clausal reduction and ellipsis is found in 27 percent of the items, and parallel structures or lists in 20 percent of the items.

In summary, we can conclude that the vocabulary portion of Section 3 tests vocabulary recognition skills as intended; other grammatical and sentence level features extending beyond vocabulary are integrated into the stimulus portions of items.

2.4.2 Section 3: Reading Comprehension. The most obvious observation about the reading comprehension subsection is that its reliance on passages as the basis for questions results in a rich sample of language for examinees to work with; stimulus passages involve almost every skill in the checklist dealing with grammatical competence and many with discourse competence. Of course, the comprehension of reading passages does not involve pronunciation recognition skills. Instead, all five passage stimuli require a large array of written symbol recognition skills. Graphic symbols used to mark formal identity occur in 63 percent of the items and options, and are more common among this item type than among any other TOEFL item type.

Noun compound lexical items, found so frequently in Section 2 and in the vocabulary portion of Section 3, occur in all reading comprehension stimulus passages, but in only five out of thirty of the corresponding question and correct answer options. The various quantifications, case relations, and semantic comparisons marked by function words occur in all five stimulus passages, with the exception of comparison and instrument or means, which occur in one stimulus passage each. Inflections of many different types occur in almost every stimulus passage, but present perfect tense markings occur in all stimulus passages, and simple past in only two out of five stimulus passages.

Particularly noticeable in reading comprehension items is the variation in sentence structure. The different complex sentence patterns identified in the skills checklist were prevalent in all five passages, with embedded structures, passives, relative clauses, and other types of subordination occurring in every passage--often repeatedly. These patterns are known to be typical of advanced expository prose, such as the prose found in college textbooks and technical writing, and, consequently, the patterns were expected. Likewise, the discourse skills required to comprehend items are quite varied and prevalent at both the lexical and sentence levels. In contrast to other item types, reading comprehension stimulus passages rely much more frequently on devices to bind interclausal semantic relationships, e.g., as realized by use of "conjunctive adverbs" serving the semantic

functions of addition, contrast, result/conclusion, time, and condition. The most frequent of these functions is the use of adverbs to denote a contrast between pieces of passage information; this device occurs in four out of the five stimulus passages for reading comprehension items.

The rhetorical and semantic organization of ideas in reading comprehension passages was not always easy to code. Many of the discourse recognition skills pertaining to structuring of ideas in written text suggested in the skills checklist--e.g., classification, cause and effect, illustration, definition, description, comparison, and chronological order--were not judged to be present in any of the passages. Two passages clearly use event ordering according to the significance of events as an organizer for a passage; the other three chain information together in no apparent sequence other than as a list or as an elaboration on a particular idea initially presented. Whether the skills checklist provided an adequate system for categorizing the structuring of planned texts typical in academic reading and writing is open to debate. The alternative possibility is that the textual stimuli presented, although authentic, were not sufficiently long or sufficiently typical to tap the various patterns of organization expected in college-level reading and writing.

Two sociolinguistic competencies deserve mention. The topics dealt with in the reading comprehension passages seem fairly typical of academic classroom content. The second observation refers to the questions and correct response options. Exactly 50 percent of the questions for reading comprehension items utilize a genre that is not common in real-life communication, limited functional segments; these are isolated segments of language that are to be interpreted as "complete me" questions by examinees. In these instances, questions are incomplete statements, completed by the correct response options to form complete sentences. As mentioned previously, even though the question and correct response option form a complete statement, adequate as a full sentence, a period is <u>never</u> part of the correct response option. For example,

Question: According to the passage, most mosquito larvae develop

Correct response: (D) in bodies of still water

Interestingly, this omission of a period is required for items of this type because of ETS item editing policy to which the TOEFL program must conform. One other observation can be made about this question type: three out of fifteen total occurrences of this genre begin with the same phrase: "According to the passage...." In fact, in examining all thirty questions in the reading comprehension subsection, 50 percent begin with this introductory phrase. The phrase does help ensure that examinees attend carefully to the passage, but one wonders whether a greater range of synonymous phrases could be used in order to diversify the language stimuli.

In conclusion, as was mentioned previously, it seems obvious that the reading comprehension portion of Section 3 draws on a great variety of discrete English language recognition skills and, even more importantly, that it requires the integration of these skills in a manner suggestive of reading in authentic academic settings. It is also notable that the content of reading comprehension items seems to require minimal knowledge of culture and life in the United States. Questions for reading comprehension items prompt the examinee to focus often on literal meaning conveyed in passages and, in particular, to main ideas and details. There are occasional requests for implicational or interpretative meaning, and in these instances the examinee must infer the writer's viewpoint, attitude, or tone. In other questions the examinee must infer details that are not explicitly stated in reading passages; in these circumstances the examinee is urged to actively guess, speculate, or infer consequences stemming from the information in passages. This prompt is indicated by the use of words such as "probably" in questions.

2.5 Concluding Comment on the Checklist Findings

Interpretation of a single TOEFL examination in terms of the competence skills checklist, as in the present study, is exploratory, and it is not intended to comprehensively investigate the competencies required by TOEFL or to evaluate it as a generic test for the assessment of English language proficiency. In reviewing the results that have been reported, it is important to ask how similar the findings would have been if other sample TOEFL tests had been investigated. A more appropriate strategy may be to ask what checklist features are shown and not shown by a collection of TOEFL tests. A further refinement of this strategy would be to study the variability and stability in competence skills required across TOEFL examinations and thus to investigate the homogeneity in coverage of skills on the test.

3. Findings Regarding Test Performance Factors Cited by Bachman and Palmer

3.1 Overview of Issues

Consistent with the discussion of Section II of this report, concerning design of the framework to describe TOEFL, the concern here is with personal, situational, and test format factors that affect the language skills used by examinees taking the test. As mentioned in Section III under "Procedures," it will only prove feasible in the context of the present study to discuss, in a preliminary way, how the various test performance factors cited earlier might influence examinee behavior on TOEFL.

3.2 Psychophysiological Skills

Language production is not required in taking the TOEFL test, and hence, psychophysiological skills related to speaking and writing are not required (finger, hand, and arm motor skills are required in marking answer options on the test answer sheet). Section 1 of the test requires aural

perception of language stimuli on all three subsections. TOEFL administration procedures, if followed appropriately, guarantee that examinees with normal hearing will have no problem clearly hearing stimuli. The fact that oral communication is not physically present as in face-to-face interaction in Parts B and C of Section 1, however, means that examinees cannot utilize visual perception to recognize proxemic and gestural cues that would enhance understanding of face-to-face interaction in real circumstances.

Because all TOEFL sections require visual processing of printed materials, acuity of vision is critical to test performance. The printed format of the test is appropriate for persons with normal vision.

3.3 Representation of Knowledge about Language and Monitoring Factors

The major issue at stake here is the kind of attention--deliberately conscious or unconscious--that examinees give to language qua language while taking the TOEFL test. If examinees consciously attend to the language of the test, calling into mind specific facts or hypotheses that they have learned as part of language instruction, they are said to be "monitoring" their language use. Such conscious examination of language has been discussed by a number of sociolinguists, and psycholinguists (e.g., by Hudson, 1980 and by McLaughlin, Rossman, and McLeod, 1983). The notion of "monitoring" referred to here is that referred to by Krashen (1982); it can only occur consciously as a deliberate effort to evaluate language as it is produced. In this case, an individual's representation of knowledge about a language is about language as an object of formal instruction. As proficiency in a language increases, monitoring of language use typically decreases. Most persons who acquire knowledge of a second language through rich exposure to it find it unusual or unnecessary to consciously monitor their second language use. There is an exception to this pattern: some persons become highly proficient in a language and yet very actively monitor their second language use. Certain linguistic structures, such as idiomatic expressions, for example, are recognized automatically by most persons who know a language well, but persons who are learning a language through classroom instruction often have to stop and recall that they have learned an idiomatic expression as a distinctive unit, prior to understanding it or producing it.

One can only speculate, of course, on how representation of knowledge of English as just described might affect TOEFL performance. It is probable that almost all TOEFL examinees engage in much conscious evaluation of the English they encounter on the test. This is to be expected since the multiple-choice format of TOEFL test items, and the motivation to pick the correct option from among distractor options, require careful evaluation of the English in stimuli, questions, and answer options. This possibility seems more likely because distractor options for test items are so often designed to incorporate features that would lead examinees with low English proficiency to consider those incorrect options.

3.4 Language Use Situations

Based on what has been suggested in the discussion of the likely role of monitoring factors in TOEFL performance, it seems evident that taking the TOEFL test corresponds more to a "reflexive" situation than to a "transitive" situation, according to Bachman and Palmer (forthcoming). Taking TOEFL is reflexive rather than transitive because the focus of the test-taking situation and performance on test items is very much on appropriateness of language itself, rather than on communication of cohesive meaning as a primary objective. Further, TOEFL meets Bachman and Palmer's definition of a "nonreciprocal" language use situation as opposed to one that is "reciprocal." Examinees are exposed to language while they take TOEFL in only a receptive manner. Thus, there is no opportunity for back-and-forth sharing of language between examinees and others during the test.

The judgment that TOEFL displays the qualities of reflexive and nonreciprocal language use situations implies that examinees' encounter with language on the test lacks the dynamic qualities typical of much everyday face-to-face interaction. While this may be viewed as a shortcoming of the test, it should be noted that the reflexive and nonreciprocal language use emphasis of TOEFL is appropriate for assessment of some language use skills of basic importance to educational performance. There are authentic language situations (e.g., reading a William Safire column, looking up a word in a dictionary, or reading a novel by James Joyce) that also are reflexive and nonreciprocal instances of language use. The critical issue would seem to be, how wide is the range of proficiency skills that one can assess based on students' encounter with language as used on the TOEFL test?

3.5 Context and Message

Four concerns are raised with regard to context and message. First, there is concern about the ratio of familiar to new information in language occurring in TOEFL. Familiar information is taken to be information that is conveyed in TOEFL test items, but that is already information known to examinees. In contrast, new information is information that examinees are exposed to for the first time when they take the test. This rendering of these two concepts seems overly simplistic, however. The author of this section of the report (Duran) did not uncover a clear perspective for interpreting the extent to which TOEFL items present examinees with familiar versus new information. A straightforward analysis of this question is inhibited for a number of reasons; the primary one is that no obvious criteria can be formulated to distinguish known from new information without possessing knowledge of this fact from the viewpoint of individual examinees.

A second issue concerns the distribution of meanings conveyed by language in the TOEFL test over time and/or space on the test. TOEFL items and questions are brief, and questions occur in close proximity to the language stimulus portion of items. Part C of Section 1, minidialogues/extended conversations, requires examinees to hold in memory an adequate recollection of an extended spoken message in order to answer three to five questions. It

is conceivable that performance on this item type could be affected by a decay in memory for an item stimulus over time. This hypothesis may not be totally appropriate, however. Psycholinguistic research (e.g., Clark & Clark, 1977) informs us that there is an interaction between memory capacity for verbal information and facility in decoding and comprehending language. Thus, persons who perform better on Part C items may do so in large part because they have greater English proficiency and, subsequently, because they are more effective in storing the information they comprehend. This hypothesis deserves empirical attention with regard to TOEFL test performance.

A third issue concerns the abstract versus concrete semantic characteristics of language occurring in the TOEFL test. Overall, the language does not appear to be unduly abstract for college-level students. Many items are concrete, referring to commonplace social roles, entities, and situations, which are everyday. Section 1 items in particular display this quality. Other items, perhaps the majority in Sections 2 and 3, have more abstract content, referring generically to concepts, classes of objects, personal and group actions, and general situations. The abstract ideas conveyed by individual sentences or by fragments of discourse on the test are relatively easy to comprehend by educated, native speakers of English. It seems clear that in order to understand the English in TOEFL, examinees must have been exposed extensively to literate varieties of English appropriate to academic contexts and to thinking in the abstract at the college or precollege level. The content of the test is probably too abstract and academically oriented for use in assessing English proficiency for a general populace.

A fourth issue concerning context and message pertains to the intensity of thinking and problem solving activity that is induced by language in the TOEFL test. TOEFL does require examinees to engage in extensive problem solving and related cognitive activity, since the goal of examinees on individual items is to identify the one correct response option. This problem-solving activity, of course, centers on properly interpreting the language of test items. Examinees need to recognize language conventions that are appropriate to multiple-choice tests but not to other forms of communication.

In looking at the presentation of TOEFL items and instructions from a discourse framework, some interesting observations can be made. In Section 1 and Section 3 (vocabulary and reading comprehension), the examinee must play the role of audience since stimuli of various textual and contextual make-up are presented, including spoken isolated statements and texts. In some cases the request for information from the examinee about the stimulus is presented in the form of detailed instructions that precede all item stimuli (e.g., statement items, Section 1; structure and written expression items, Section 2; vocabulary items, Section 3), but in other cases the request for information about the stimulus is given after its presentation (e.g., dialogue and minitalk/extended conversation items, Section 1; reading comprehension items, Section 3). When the request for information about the stimulus follows the stimulus, it is often in a complete interrogative form (dialogue items and minitalk/extended conversation items, Section 1, and some reading comprehension items, Section 3). However, it may also frequently occur as an

incomplete statement, as is typical of many reading comprehension items in Section 3. Thus, in light of the range of variation in presentation of items and item types, an answer option for a TOEFL question may complete a communicative event initiated immediately preceding it (i.e., a question and answer text may be formed). This is true of dialogue and minitalk/extended conversation items, for example. In contrast, an answer option may complete a communicative event initiated far in advance through a series of instructions about how to take a TOEFL test section.

There are numerous instances throughout the test in which some aspects of natural discourse are simulated through the question-option context typical of natural dialogues. The most obvious display of this sort of authenticity is the use of elliptical statements offered as options in over half of the dialogue and minitalk/extended conversation items of Section 1.

Finally, it is useful to offer a comment concerning the extent to which the TOEFL test draws on inferential skills or conclusions based on material not supplied directly in the stimuli. To correctly answer some TOEFL items, inferences are required that reflect extended knowledge of semantic conventions in a language and cultural norms. Some examples of inferences that examinees must make include:

(a) if one is "behind," then consequently they have to "catch up"

(b) if the seminar "broke up," this means everyone then "left"

(c) if a gentleman has his "hands full," he will accept help from a female in carrying the load

(d) reference to "front tire and seat lowering" are applicable to a "bicycle"

(e) going "straight" to a place refers to time not to spatial orientation.

3.6 Artificial Restrictions

The concern here is with restrictions in language performance that arise from the format of TOEFL test items. The focus is on the qualities of the language samples that examinees encounter on the test and whether the language in the test as a whole has authenticity in comparison to language arising in criterion settings. This question is to a large extent covered in the previous discussion on context and message, and by the discussion in Section 2 of the findings concerning the competencies required in performing on the test. The question is also briefly addressed in the next section of the findings, concerning the social and academic naturalness of language occurring in TOEFL. There is no doubt that language in TOEFL manifests artificial restrictions that result from the format of test items. It should be noted, however, that there are testing program operational restrictions in

the design and use of specific test questions to assess discrete language proficiency skills. Discussion in the last section of this report will review the implications of this conclusion for the TOEFL test and for other tests of language proficiency.

3.7 Affective Factors

Affective responses of examinees to TOEFL test items is an unexplored matter. The issue, as presented by Bachman and Palmer (forthcoming) concerns the affective orientation, cognitive style, and motivation that examinees show while performing on the test. Certainly, one key matter involves the motivation of examinees to perform well, given that TOEFL scores may be a critical factor in the students' admission into a North American college. The central question is whether there are affective factors that can enhance or degrade performance on a test such as TOEFL for some individuals. At present we do not have empirical evidence on this question, and it would seem valuable to gather such information. Two practical goals for such research stand out. First, can we learn how to reduce stress, anxiety, and other affective factors by further counseling prospective test takers on good test-taking strategies? Second, are there new test item formats that can reduce the negative performance impact of affective factors? Canale (1984, in press and in part of Section V of this report) suggests that there are ways in which both of these goals can be met. For example, design of multiple-choice proficiency tests might be altered so as to allow examinees to perceive a more naturalistic and cohesive flow of information and questions. This tactic might enhance examinees' ability to demonstrate their language competencies more fully and unobtrusively on language proficiency tests.

3.8 Strategic Factors

Since TOEFL requires examinees to passively understand, but not produce, language, they have no opportunity to interact with others or to negotiate meaning. It is also interesting to note that Section 1 of the test does not noticeably feature strategic competence on the part of interlocutors in the stimulus portions of items. One can infer that the absence of strategic competence is due to the test development goal of making Section 1 item stimuli maximally comprehensible and brief.

4. Findings Regarding the Relevance of TOEFL Item Content for Academic and Social Language Use

4.1 Rating Agreement

Items from the sample TOEFL test were each rated on three dimensions of representativeness: academic living, academic content topic, and social naturalness (the last scale was used only for Section 1 items). A rating of 1, 2 or 3 was assigned each item on each scale. A 3 rating represented a judgment of clear and direct authenticity, while a 2 represented a judgment of less clear, but possible authenticity. A 1 rating indicated a judgment of ambiguous or low authenticity. Rating data were generated by two independent

coders. The data that resulted for the criterion coder are summarized in Section B 1.2 of the checklist that composes Appendix D. The entries are mean rating scores for item types within given sections. Mean ratings are also provided for item types collapsed within a section.

Agreement of the independent coders was assessed by computing the number of identical ratings for items. Evaluation of the degree of agreement among raters is important, since achieving a reasonable level of agreement is necessary in order to have confidence in the findings to be discussed. Agreement was lowest for Section 2 of TOEFL and for one item type in Section 1. On Section 2, the raters agreed on only about one-half of the judgments they made about the academic life representativeness or academic content of items. This level of disagreement is due to the fact that one rater rated more items at a 1 level more often than did the other rater. Agreement levels were in the range of 66 to 70 percent on Section 1 for the academic living and academic content representativeness scales. For Section 1, there was only a 48 percent agreement in ratings with regard to the social representativeness scale. Agreement rates ranged from 77 percent to 87 percent across the academic living and academic content representativeness scales for Section 3. The low agreement rates are indicative of the difficulties encountered in providing raters with adequate criteria and training on which to base scale judgments because of the limited information about the pragmatic context alluded to by the content of items. Thus, judgments of the representativeness of language on items on each scale had an ad hoc character. As the ensuing discussion will show, however, some useful conclusions can be drawn from the data. In defense of these conclusions, it should be pointed out that raters' judgments tended to cluster and that they disagreed more in making ratings of 1 versus 2 than in making ratings of 2 versus 3. In the course of making ratings, judges only differed once by 2 scale points.

4.2 Ratings of the Characteristics of Section 1 Items

As shown by the data in Section B 1.2.1 of Appendix D, the criterion judge found that Section 1 TOEFL items had relatively high social authenticity, but lower academic living and academic content authenticity. The high social authenticity ratings of Section 1 items must be interpreted cautiously, because they reflect a judgment that language clearly arose from social interaction, but not a rating of how much information there was about the nature and course of the social interaction. Despite the high social authenticity ratings, there seems to be no doubt, from a communicative viewpoint, that the social contexts depicted by TOEFL Section 1 items are limited in their rendition of everyday social contexts and communicative competencies exercised in everyday social life. With the exception of minitalk/extended conversation items, ratings of the academic living representativeness of Section 1 items averaged below 2. Minitalk/extended conversation items were judged to involve language that was more representative of situations encountered in academic life than did statements and dialogue items. Ratings of academic content representatives of topics brought out in Section 1 TOEFL items were low for statement and dialogue items, as one might expect. Minitalk/extended conversation items were found to involve topics more closely related to academic content.

4.3 Ratings of the Characteristics of Section 2 Items

Structure and written expression items averaged a 1.5 rating in their representativeness of academic living and academic content. No 3 level judgments were made for these items on either scale. Thus, about half of the time, the criterion rater judged that items were ambiguous or not representative of academic life situations or academic content material. The extreme brevity of items in Section 2 is, no doubt, responsible for this pattern of ratings; items do not have enough linguistic context and situational reference to show clear authenticity with regard to academic living situations or academic content topics. Interestingly, and in contrast to this finding, all the materials used in Section 2 are based on sentences taken from encyclopedic sources or textbooks. The extraction of sentences from the original texts appears to result in a reduction of the academic authenticity of item stimuli when such authenticity is determined by our criteria. While Section 2 items may be useful for assessing discrete grammatical and lexical skills, they are not examples of the extended-length texts that foreign students encounter in college life and college academic study areas, outside of the confines of similar multiple-choice language proficiency tests.

4.4 Ratings of the Characteristics of Section 3 Items

Vocabulary and reading comprehension items averaged 2 or a little under 2 in their judged relevance to academic living situations or presentation of potential academic content material. Based on the definitions of the ratings, the overall conclusion to be drawn is that the criterion rater judged, on the average, that vocabulary and reading comprehension items might be relevant to academic life or to college-level academic content materials, but that there was no clear and compelling evidence to assert overwhelmingly that they were on the average. Only two items in all of Section 3 were judged to involve language that unquestionably might come from an academic living situation or that was exemplary of college-level academic content materials. The ratings, as a whole, reflect the absence of information concerning the pragmatic meaning that could be attached to the content meaning of items.

In passing, it should be noted that the criterion coder found that some of the Section 3 items seemed more representative of high school subject matter than of college subject matter.

5. ILR Scale Description of TOEFL Proficiency Levels

This section describes use of the ILR scale to rate the proficiency level required to answer items on various TOEFL sections.

There is some empirical evidence that the ILR scale is relevant to performance on TOEFL. In a study involving thirty-one foreign teaching assistants at American universities, Clark and Swinton (1980) found a moderate relationship between TOEFL scores and rating on an oral proficiency interview employing a scale similar to the current ILR scale. The correlation between scores on Section 1 of TOEFL and ratings on the oral proficiency scale was

.71; for Section 2 it was .57; for Section 3 it was .62; and for TOEFL total score the relationship was .71. This evidence must be interpreted with caution, however, since it may reflect the relationship between TOEFL scores and examinee performance, rather than between TOEFL scores and the situational/linguistic input material that is presented to the examinee. Nonetheless, the situational/linguistic input does serve as the context for examinee performance, and as mentioned earlier (see page 23) situational and linguistic input are organized hierarchically and therefore form a part of the descriptions of the various levels on the scale. Thus, it seems reasonable to describe the content of a second language test using the ILR scale.

Before discussing the findings, it should be noted that the ILR scale may be more suited to the analysis of integrative rather than discrete-point tests of language proficiency. This is because one criterion for a rating of level 2 and above is the availability of a corpus of language equivalent to a paragraph or more in length for any given communicative situation. Since discrete-point items are usually limited to one sentence of text or less, one encounters a diminished situational context. Thus, even though the linguistic complexity of the language in the item may be indicative of higher level skills, most discrete-point items must be classified below level 2, that is, at level 1+ or below.

5.1 Section 1: Listening Comprehension

Part A of the listening comprehension section is typical of this dilemma. This part of the test contains twenty restatement items in which the task is the literal comprehension of a single sentence that deals with concrete facts. All items in this section represent listening proficiency levels 1 or 1+, with the majority falling into the latter category. Listening proficiency level 2 involves the comprehension of multiple facts in an extended text. In Part A of this section, the message is limited to a single sentence, and the options focus on comprehension of the main idea or a single fact contained in the stimulus. Thus, most items were rated a level 1+. Only a few level 1 items were found, and these came at the beginning or the middle of this part.

The first item in Part A is illustrative of a level 1 item.

*(A) Go directly to the post office when class is over.
(B) Let's first straighten up the classroom and then go to the post office.
(C) That's the most direct way to the post office from our class.
(D) The post office is straight ahead of the classroom building.

The task involves understanding simple directions one sentence long. The directions relate to "getting around," which is a "survival" (i.e., level 1) task. The stimulus is not complex; it is a simple sentence using an imperative form of the verb "to go" and everyday vocabulary. Studies in second

language acquisition have found that the imperative form of English verbs is usually acquired prior to the English present tense (Burt et al., 1976).

Item 3 in Part A illustrates a level 1+ item and is typical of the general level of difficulty of items in this part.

Greg thought he could do it himself.

*(A) Greg believed he could do it alone.
(B) Greg thought he'd cut himself
(C) Greg thought he was selfish.
(D) Greg alone believed it could be done.

The stimulus involves an independent clause in the past tense followed by a dependent clause in the conditional tense. The dependent clause also contains a reflexive pronoun, which is the key point tested in the item. The topic is not a survival topic, but instead makes mention of a routine event that might be the subject of a social encounter. Thus, while the syntactic and situational characteristics of the sentence are at level 2, the item must be classified at level 1+ on the ILR scale due to insufficient context.

Some stimuli in this part contain words or idioms that also exceed level 1+ in difficulty, such as item 19.

The encylopedias were out of order.

The subject is a low-frequency word that could pose a problem for lower and intermediate level learners whose native language does not use a cognate form. The usage of "out of order," meaning improperly arranged, can be confused with another usage of the same idiom, meaning out of service. Indeed, one of the distractors focuses on this type of semantic interference. Because only six words are involved, this would also be classified at level 1+ on the ILR scale. The classification illustrates how the ILR scale can underestimate the communicative competence required to handle discrete-point items.

Part B. This part of Section 1 consists of short conversations involving either a man and a woman or two men. At the end of each conversation, a third voice asks a question about what was said. Each speaker in the conversation has one communicative turn. Thus, the conversation involves a single exchange of messages rather than multiple exchanges.

Since the conversation involves more than a single utterance, this item format permits the design of items that are at level 2. Nevertheless, few of the items in this part test level 2 listening skills as defined on the ILR scale. This is because the conversations are normally short, followed by a single question that usually requires comprehension of only one utterance, instead of both. (See example items 22 and 35 on the following pages.) Since the comprehension of a single utterance is a level 1+ task, such items are lower than level 2 in difficulty. In fact, most of the short conversation

items examined were classified at level 1+, while some were at level 1, and a few were at level 2.

Item 21 illustrates a level 1 item.

(man) Good morning, may I help you?

(woman) Yes. I'd like to cash these traveler's checks first and then open a savings account.

(voice) Where does this conversation probably take place?

(A) In a department store.
*(B) In a bank.
(C) At a tourist bureau.
(D) At a hotel.

The item depicts a survival situation similar to one that might be encountered by a tourist on a short sojourn in another country. The examinee need only comprehend the general nature of the message by indicating where the situation takes place; the item does not require the comprehension of detail. The options are phrases rather than sentences. Thus, they display the same ellipsis that is typical of normal speech. While all other Part B items on this test form contain full sentence options, the options are often slightly shorter than those found in Part A.

Although the stimulus conversation depicts a simple survival situation, the question invokes processes that are representative of normal conversation, i.e., a reliance on shared knowledge and inference. Such processes are typical of natural discourse.

As indicated above, the short duration of the stimuli often prevents items from testing level 2 skills. Although no long stimuli were found in this form, some do occur in other editions of the test. Item 22 in Understanding TOEFL: Test Kit 1 exemplifies such items.

(man) I can't understand why my friend isn't here yet. We agreed to meet at 10:30. It's almost 11:00. Do you think we should try to call her or go look for her?

(woman) She probably just got tied up in traffic. Let's give her a few more minutes.

(voice) What are these people going to do?

(A) Check the time of high tide.
(B) Go stand under the clock.
*(C) Wait a little longer.
(D) Look for the traffic light.

In this item, both utterances contain more than one sentence. The conversation meets routine social demands, and the question calls for an inference.

On the form used in this study, item 35 can also be classified at level 2.

(man) I heard that the newspaper gave that book a terrible review.

(woman) It depends on which newspaper you read.

(voice) What does the woman mean?

(A) You should believe everything you read.
(B) She thinks the book is excellent.
(C) She wonders which newspaper he reads.
*(D) Reaction to the book has been varied.

The topic of this conversation is clearly above the survival level that is associated with level 1 of the scale. Also, the ability to answer the question depends on comprehension of both utterances. Finally, in order to answer the question, one must make an inference, since the correct response is not actually stated during the conversation. This ability to "understand between the lines" is associated with level 2+ on the scale. Nevertheless, the short duration of the stimulus means that the item cannot be classified beyond level 2 on the ILR scale.

In summary, most items in this part were found to fall at level 1+. This is due to several criteria that are part of the ILR scale: i.e., the items contain short stimuli; the topic or situation is often at the survival level; the examinee is often required to understand only a single utterance.

Part C. In this form of TOEFL, Part C consists of three listening comprehension passages. Two of them are lengthy, one is short.

The first passage is a monologue or short lecture about the high rate of deafness on the island of Martha's Vineyard, Massachusetts. (See script for the listening comprehension section, Appendix A, page 8.) The passage contains seven sentences and is just over two hundred words in length; it can be classified at level 3. Although the monologue is not technical, it is a narrative that describes a sequence of situations and events and uses a deliberative style (Joos, 1967). In this sense, it is comparable to a newspaper article that discusses a topic of general interest. (Such articles are considered to be at level 3 on the ILR reading scale.) The same type of material might be heard on an educational radio program. The passage requires the examinee to follow a series of facts and then put together ideas that are interwoven with these facts regarding genetic breeding, the dominance of certain defects, etc. While the passage is definitely at level 3 in difficulty, many of the items test listening ability at level 2. They focus on facts, rather than on inferences and conclusions that can be drawn from the passage.

The items could be classified at level 3 if phrased differently. An example is item 41, which is based on the last sentence in the passage.

> (last sentence) In the twentieth century the local population has mixed with people off the island, and the rate of deafness has fallen.
>
> (voice) According to the talk, how has the island changed in the twentieth century?
>
> *(A) The patterns of marriage have changed.
> (B) Many deaf people have regained their hearing.
> (C) Most of the original population has left the island.
> (D) The island has become famous for its research facilities.

The question posed by the voice calls the examinees' attention to the time frame (twentieth century) in which the action in the key occurred. Thus, no inference is required. On the other hand, an inference is signaled by the conjunction "and" in the last clause of the passage. The ability to make this inference could be tested by a similar question, such as "The rate of deafness on the island has decreased recently because." Such a question focuses on the type of listening task (inferring or concluding) that is often posed to level 3 learners of a second language. Still, this passage is the most demanding communicative task on the listening comprehension section when analyzed according to the ILR scale. It is also representative of the type of task required of a student listening to a lecture in a university classroom.

The second passage in Part C is an extended conversation between a man and a woman about California redwood trees. Each person speaks five times. Thus, the conversation contains a total of ten turns. While a considerable amount of information is divulged regarding redwood trees, the passage is not as technical as the monologue. Also, its conversational tone of a question followed by an answer tends to diminish the information load by separating it into segments. As a result, the examinee only has to listen to the information; it is not necessary to piece together a sequence of facts so they take on a larger meaning. The passage requires less cognitive processing than the monologue. The only sequence that one has to process is that trees with thick bark that live in a damp, foggy environment do not catch on fire, and therefore live a long time. Since the circumstances are contiguous in time and the conclusion is logical, the passage does not invoke the same degree of proficiency as the previous passage, which explains that repeated intermarriage produces a special set of genetic circumstances. Thus, when deafness has occurred in one generation, it may also show up several generations later because the propensity for it is carried by both parents. Because the conversation passage is easier to process than the short lecture, it would be classified at a 2+ on the ILR scale.

The items following the conversation are also straightforward. Most are restatements of the questions asked by one of the participants, and the answers are based on the responses given by the other participant. Thus, the examinee's attention is called to key facts by questions that are part of the stimulus.

The final listening passage is a short announcement of the closing of a library at the end of the day, similar to one that might be heard over a public address system. Such announcements about daily events are classified at level 2 on the ILR listening scale unless they are presented under adverse conditions, such as a high amount of static noise, etc.. Radio news broadcasts, on the other hand, are usually classified at level 3 because the listener must comprehend not only the facts, but also analyses and interpretations made by the commentators. The questions deal with the factual content of the announcement and thus are designed to test level 2 listening skills.

5.2 Section 2: Structure and Written Expression

This section of the test is composed exclusively of one-sentence discrete-point items whose function is to assess accuracy of grammar and style. In a general sense the observation is warranted that these items do not serve a communicative purpose. Thus, they are not well suited for analysis using the ILR scale. Nevertheless, some judgments can be made about the degree to which they test communicative competence using certain points raised by the scale as a criterion.

The section consists of two parts. Part A, structure, consists of fifteen incomplete sentences. In each sentence a word or phrase has been left blank. The examinee is to select from four options the one that best completes the sentence. The items focus on a wide variety of grammatical problems. Part B consists of twenty-five items in which four words or phrases in a sentence have been underlined. The examinee must identify the one underlined word or phrase that needs to be corrected or rewritten.

All of the sentences in this section are appropriate for level 2 texts or higher. One can imagine many of these sentences appearing in textbooks. Thus, lexically and structurally they are at level 3, although there is insufficient context to classify them as such. For the items to be more communicative, it would be necessary to relate them to each other or to place them in a text. A multiple-choice cloze is an example of such a context, although cloze passages usually emphasize lexical and grammatical problems equally.

Structure items seem to be slightly more communicative than written expression items. Generally, they are slightly longer and thus offer more context. In general, additional context can be used to assess more long-range constraints on an item. Many structure items function like a cloze exercise in that they utilize the constraints provided by the context of the sentence to help one select the correct completion. Written expression items focus only on the recognition of syntactic and stylistic problems.

5.3 Section 3: Reading Comprehension and Vocabulary

This section is designed to measure the examinee's ability to understand various kinds of reading materials, as well as the ability to understand the meanings and uses of words. It consists of thirty vocabulary items and thirty reading comprehension items based on reading passages. The length of the passages ranges from one hundred to three hundred and fifty words.

In the vocabulary items, the examinee is presented with a sentence that has one underlined word. The task is to select which of four options best keeps the meaning of the underlined word. From the standpoint of the ILR scale, the vocabulary items focus less on communication of meaning in context than do any other parts of the test. This is so despite the fact that examinees would need to exercise grammatical recognition skills in order to understand stimulus sentences. At least half the items can be answered without reference to their context within the sentence. Item 1 is typical of such items.

> In masculine rhyme, the end sounds of stressed syllables are repeated.
>
> (A) dominant
> (B) vowel
> (C) hard
> *(D) final

While the options are all adjectives that might be associated with the noun "sounds," and therefore effective distractors for an examinee who does not know the meaning of the underlined adjective, it is not necessary to understand the context in order to answer the item correctly. This is not true of all vocabulary items, however. The following item (item 12) is context sensitive. That is, only by reading and understanding the whole sentence can one identify the correct answer. In a different context, any of the options might be considered a correct synonym for the underlined word.

> The Cheyenne Indians were considered spectacular riders and fierce warriors.
>
> (A) hunters
> (B) defenders
> (C) enemies
> *(D) fighters

The vocabulary items tested in this part vary in difficulty. They would be part of the receptive lexicon of readers at level 2 to level 4. Within this range some lower-level items are end, quarter, shy, and materials. Some items that would be part of a higher level lexicon are heyday, incessantly, ban, and hailed. The nouns in this test form generally are easier than words belonging to other parts of speech, although the actual difficulty of items for each examinee would depend in part on whether the examinee's native language contained a word that shared the same derivational morpheme.

Reading Comprehension. Each form of the TOEFL test contains thirty reading comprehension items, although the number of passages may vary from form to form. The length of passages varies also, from one hundred to three hundred and fifty words. On this form, there are five passages ranging from 103 to 263 words in length. Each passage is followed by four to eight questions.

The five texts vary in difficulty from level 3 to level 4. The first passage is at level 4; the second and third are at level 3; the fourth is at level 3+; and the final passage is at level 3. All are the type of passage that might be encountered by a university student at the undergraduate level. Thus, reading comprehension appears to be the most demanding portion of the test in terms of communicative competence when judged by the ILR scale.

The first passage deals with mosquito larvae and is the most demanding. There is a heavy information load, and the passage is lacking in cohesion, characteristics that typify many authentic texts. Although the passage discusses a scientific topic, it is written for the general public rather than for scientists. One of the characteristics of a level 4 reader is the ability to "follow unpredictable turns of thought." The fact that this passage was taken from a longer text, and discusses a technical topic about which few members of the general public would have knowledge, suggests that it would invoke such skills. Another characteristic of an ILR level 4 reader is the ability to recognize all professionally relevant vocabulary known to the educated nonprofessional native. Again, the passage appears to invoke this skill.

The second passage, on oranges, is at level 3. It is a factual and historical passage written in a formal style.

The third passage, which deals with communication satellites, is also technical in nature. It employs a higher level of vocabulary, including one instance of communications industry terminology ("domsats"), which is defined in the passage. Because of its vocabulary, it might be classified as a level 3+ text, although in practice a government examiner would not consider such a short passage (103 words) to be at the 3+ level. Because the passage is not sufficiently long, it must be classified at level 3.

The fourth passage, which describes a particular type of animal behavior, contains a few technical terms along with many words that are often known by only advanced users of a second language and educated native speakers. Some of these words are sensory, vertebrates, diurnal, predators, flock, mammals, rodents, herds, hoofed, prey, moose, dens, primate, and baboons. In spite of its advanced vocabulary, the passage makes frequent use of cohesive devices, is well organized, and could be understood by a level 3+ reader. The passage is also the longest on this form (263 words).

The final passage, which discusses the problems of doing good research, is also a level 3 text. Although the vocabulary is suitable for a college audience, it does not contain low-frequency words that exemplify the highest levels on the ILR scale. While the topic discusses an academic

situation, the text is a narrative and relates a frequently encountered personal experience.

The items following each passage are appropriate to level 3 texts. Quite often the questions seem to avoid testing comprehension of less frequently used terms that appear in the text, unless these are specifically defined. For instance, the passage on animal behavior is followed by questions that generally avoid the vocabulary listed above, and test comprehension of content that is also unrelated to the phrases or sentences in which this vocabulary appeared. Thus, even the more difficult reading passages do not seem particularly suited to assess highly advanced (level 4) users of English as a second language. This may not be inappropriate however, since a test designed for such bilinguals would probably confuse many native English speakers with nonnative speakers, which is a problem the current TOEFL does not have.

5.4 Conclusion

The communicative competence invoked by TOEFL items seems to be quite consistent within each item type, while deviating considerably across item types. In general, the discrete-point items were rated low on the ILR scale (from level 1 to level 2). Of these, vocabulary items generally made the least use of situational context. The monologues and extended dialogues in Section 1 and the reading passages in Section 3 were found to be the most demanding communicative tasks on the test; text difficulty ranged from level 2+ to level 4. The test items were frequently found to be of a lower level than the passages on which they were based. Even with the more difficult listening and reading passages, the more advanced communicative tasks were not assessed by the comprehension questions.

On the basis of this analysis of the test using the ILR scale, it appears that the communicative content of TOEFL, while limited in some ways, is appropriate for assessing the language proficiency of basic, intermediate, and advanced learners and of fluent users of English as a second language. However, it does not appear that the test is appropriate for identifying or discriminating among highly proficient, near-native speakers. The discourse tasks tested are not that advanced, nor are they as communicatively rich as the range of tasks on the ILR scale. This does not mean however, that examinees who attain high scores on TOEFL (i.e., 630 and above) could not perform such communicative tasks; it merely indicates that such tasks are not measured to a significant degree on the test. To be more specific, such tasks (as defined on the ILR scale) as the ability to understand fully all input on any subject within one's range of experience and to grasp all subtleties and nuances, and the ability to understand speech under unfavorable conditions, including nonstandard dialects and slang, as well as cultural references and extremely abstract prose, are not measured by the test.

Finally, it should be remembered that the ILR scale is designed to provide a global rating of communicative competence. In this analysis, it was found that discrete-point tests that focus on linguistic accuracy are not well suited for analysis by this scale. Quite often, the difficulty of such items

is underestimated by the scale. While adherence to the requirement for a specific amount of context may appear to be unnecessarily strict, it must be remembered that the ILR scale is basically noncompensatory. Thus, the satisfaction of some but not all of the requirements for a particular classification is not sufficient reason for an examinee or a passage to receive that classification.

V. DISCUSSION AND DIRECTIONS FOR ITEM AND TEST DEVELOPMENT

1. Overview of This Section

This section of the report first discusses the major findings of the study. It then outlines some considerations that should be made in improving the design of language tests in general, and finally, the implications of some of these suggestions for the TOEFL test.

2. Conclusions Regarding TOEFL

The description of the TOEFL test given in this report suggests that the content of the test is directly appropriate for assessing some language skills but not others. This conclusion is based on a communicative description of the content characteristics of items only. No comprehensive research has been done investigating skills that are not assessed by the test but are important to foreign students' academic functioning. Research is needed on this question and on TOEFL's correlation with these skills. The test requires only receptive language skills on the part of examinees, and in this regard the main strength of the test lies in its assessment of examinees' ability to recognize orthographic encoding of information, grammatical appropriateness at the spoken or written sentence level, and recognition of spoken or written semantic information at the sentence or brief text level. As outlined in the TOEFL test specifications, each section and item type within a section are designed to assess examinees' recognition of a particular point or skill. The discrete skill emphasis of the instrument is enhanced by the frequent use of many items that are tailored to the skill being tested; with the exception of minitalk/extended conversation items in Section 1 and the reading comprehension items in Section 3, adjacent items never share contents that refer to the same topics.

Within the range of skills TOEFL can assess, the naturalness of language that examinees are exposed to is a direct function of the length and rhetorical complexity of the language stimuli that accompany test questions. For example, the minitalk/extended conversation items of Section 1 and the reading comprehension passages in Section 3 provide the richest integration of language phenomena, and, accordingly, these items exercise examinees' language recognition skills in the richest and most face valid manner. Both types of items are clustered into small groups of questions regarding an extended text stimuli. However, there is no evidence of a deliberate thematic connection in the topics pursued across questions in the same set.

The statements and dialogue items of Section 1 provide examinees with less language than do the two above-mentioned item types; dialogue items do require examinees to recognize the relation between a pair of utterances made by two speakers, but correct response options seem most related to the second utterance. A number of dialogue items feature speech act functions that go beyond simply giving and asking for information, and while this frequency of

occurrence is not rare, occurrence of speech act forms must be inferred. The nature of the social context surrounding occurrence of dialogue items must be inferred by examinees, and there is only limited information in the linguistic and paralinguistic content of dialogues to guide this form of inference. The use of paralinguistic cues, such as stress, and modulation of intonation in speech sometimes seems to be more enunciated and hyperbolic than that encountered in casual face-to-face conversation. This finding should be weighed against the need for high acoustic quality in a tape-recorded section. Delivery of dialogue items by tape recording seems to be a constraint, in that examinees are not physically exposed to gestural and other proxemic cues that can enhance inference of meaning and intent in speakers' utterances. The speech in dialogue items is devoid of false starts, hesitations, back channel cues, and other phenomena that are associated with the exhibition and recognition of strategic competence.

The statement item stimuli are the most decontextualized examples of natural speech found on the TOEFL test. These stimuli are provided with no information about the communicative context within which they occur, and while they feature intonation, stress, and prosody that are typical of North American English speakers and have referential meaning, the specific function and purpose of utterances is not always easily inferred. While the correct, written response option for statement items is typically a paraphrase of the stimulus, on occasion examinees must infer a correct response option based on pragmatic knowledge of the real world situation the stimulus and response option are judged to refer to.

Section 2 of TOEFL, structure and written expression (along with the vocabulary portion of Section 3), utilizes the most restricted and least natural written language stimuli occurring in TOEFL. The two structure and written expression item types are designed to probe examinees' knowledge of sentence grammar and appropriate usage of words, but not range of vocabulary. The two item types consist of isolated sentence fragments or else one or two isolated sentences. No information is provided about a communicative context, and the topics of different items are always unrelated to each other. The TOEFL test specifications for Section 2 suggest that this variability is a desired goal, in order to expose examinees to a greater range of contents and topics.

A review of test method and personal factors that might affect TOEFL identified a number of ways that examinees' performance might be distorted or limited relative to situations in which authentic communication occurs. All language proficiency tests, be they direct or indirect, or discrete-point oriented versus integrative or holistic, entail some distortion of normal communicative processes from the viewpoint of examinees. The performance constraints occurring on TOEFL relative to normal communication limit the assessment of an extended range of communication skills because the test involves only language reception on the part of examinees, and because multiple-choice items present diminished situational context. However, it should be noted that TOEFL is not designed to be a test of extended communication skills. Also, such a criticism is not unique to the TOEFL test; the same criticism could be made of other proficiency tests currently in use for

the same purpose as TOEFL. If the sole purpose of TOEFL is to challenge examinees with questions on specific points related to American English grammar, vocabulary, and usage, then TOEFL items perform this task. The immediate critical issue seems to be whether TOEFL might achieve its goals better with items that have greater face and content validity.

The analysis of test method and personal factors potentially affecting TOEFL performance that was presented in this report is quite preliminary. The Bachman and Palmer (forthcoming) outline of various factors and related issues has important, substantive value in describing performance constraints on the TOEFL test. However, a more intensive and careful effort is needed in order to interpret fully the underlying issues in a more definitive and analytically useful fashion than is possible in this report.

Evaluation of the content of TOEFL items with regard to their academic and social naturalness leads to the conclusion that items are limited in how well they simulate authentic language situations that may be faced by examinees when attending North American colleges. This conclusion is to be expected given that individual test items and item stimuli are very abbreviated in comparison to language use in authentic communication contexts. The absence of information about a language use context, the inability of examinees to exchange language reciprocally, and the brevity of discourse-length texts are the most critical factors inducing judged lack of authenticity of TOEFL items. The item types showing the greatest authenticity—minitalks/extended conversations and reading comprehension—also involve the grammatically, lexically, and rhetorically richest discourse of all TOEFL items. Even these item types, however, can be judged as having limited authenticity because they retain an isolated character that renders them as fragments of extended discourse drawn in an ad hoc fashion from the range of all possible social-academic situations and academic content experiences that students might encounter in college. Strengthening the naturalness of language found on these TOEFL items by revising item content specifications is possible in principle, though there are a number of practical, psychometric, and operational issues that would mitigate this suggestion. This matter is addressed in the final section of the report.

Evaluation of the minimal ILR proficiency level required to solve items on various TOEFL sections led to the conclusion that the test is appropriate in difficulty for basic, intermediate, and advanced learners of English. Taken collectively, test items are capable of discriminating persons who might be rated at all but the highest ILR levels in a given modality of language use. The upper levels on the ILR scale for a given modality involve examinees' ability to produce or comprehend language with the same fluency, cultural appropriateness, and cognitive problem-solving ability as educated native speakers. Assessment of proficiency at these levels is not the purpose of the TOEFL test, but some items nonetheless are capable of assessing the English skills of highly proficient, near-native English speakers.

Considerable variation in the ILR proficiency level exhibited by different item types was noted. In general, consistent with other findings in this report, the minimal ILR level required to work item types correctly was

directly related to the extent of discourse presented by an item type. The reading comprehension section of TOEFL seemed to consistently require the highest ILR proficiency levels.

The analysis of TOEFL that is summarized here at times might be interpreted as having a critical tone that may be misleading; research on the test has clearly demonstrated its value as an instrument for assessing the language proficiency of incoming foreign college students. The principal concern here is with understanding better the content nature of TOEFL items, and thereby some of the strengths and limits of the test in assessing language proficiency skills. A secondary goal was to suggest how TOEFL and other tests might be made more responsive to improved assessment needs that theoretical developments and research in applied linguistics permit us to describe. The weaknesses of TOEFL are not unique to this proficiency instrument, and the extensive review of the test from a communicative proficiency viewpoint provided here is unavailable currently for any similar commercial instrument. There is no reason to believe that any other similar, currently available language proficiency test would be evaluated any more favorably than the TOEFL test. However, a number of efforts are underway to develop proficiency tests based on a communicative approach to language testing (Wesche, 1981; Canale, forthcoming). Communicative approaches to language proficiency assessment have raised new challenges to language test developers. The next section of the report discusses these challenges and ways in which they might be resolved in a manner leading to improved test use and benefits to examinees and institutions. Following this discussion, attention returns to TOEFL and to efforts that the TOEFL program could undertake to develop new item types and tests, and efforts that also would strengthen appropriate use of the existing TOEFL.

3. Suggestions for Improving Language Testing

This section summarizes some recent views on general test design that suggest how language testing might be improved. First, there is a discussion of two guiding principles for improved test design. Next, an outline of four test design features consistent with these principles is presented.

3.1 Two Guiding Principles

Most discussions of test design emphasize that a good test is characterized minimally by validity, reliability, and practicality. Many major language tests do seem to reflect careful attention to content and concurrent validity, to internal and test-retest reliability, and to practicality of test administration and scoring. Two further considerations, not always emphasized in test design or reflected well in many major language tests, are test acceptability and feedback potential. Acceptability, similar to the notion of face validity, concerns the extent to which a test task is accepted as fair, important, and interesting by both examinees and test users. Feedback potential concerns the extent to which a test task rewards both examinees and test users with clear, rich, relevant, and generalizable information. Both may be hidden in a test score.

While these two considerations may be acknowledged by many test developers, they are often accorded lower priority than are concurrent validity, reliability, and practicality concerns. As a result, one risks developing and tolerating language tests that may measure something well but be neither acceptable nor rewarding to examinees and test users. A rather extreme example, noted by Clark (1972) would be that a test of oral interaction skills that consisted of pencil-and-paper multiple-choice items that never required the examinee to face an interlocutor or to speak. No doubt such a test could be validated and used. However, this would be to ignore--almost to degrade--the interests of the examinee on the one hand, and, on the other, of the educator seeking to understand and improve the examinee's skills. While such testing practices may be more defensible when large numbers of examinees are involved, they unfortunately risk becoming the model for testing small numbers of examinees as well (e.g., in classroom testing).

To counter the risk of emphasizing only psychometric and practical concerns in test development, the following guiding principles are offered.

> Principle 1: attempt to elicit the best performance from examinees by presenting tasks that are fair, important in themselves, and interesting in themselves. In Swain's (1982) terms, bias for best.
>
> Principle 2: attempt to provide test tasks that reveal to examinees and educators clear, rich, relevant, and generalizable information--in our terms, bias for rewarding feedback.

These two principles are highlighted here not because they are the only ones worth considering but because they are often not represented adequately in language tests. Furthermore, these principles are certainly compatible with suggestions calling for close attention to the variety and quality of performance elicited by language tests as well as to the naturalness of tasks on such tests.

3.2 Four Test Design Features

In keeping with these last comments, four general design features are worth considering to improve receptive language testing.

(a) Thematic organization. In contrast to a language test that might be organized according to linguistic criteria (e.g., a vocabulary section, a sentence structure section), skill areas (e.g., a sound discrimination section, a reading comprehension section), or other criteria, a thematically organized test would represent and group those tasks that provide a coherent, natural, and motivating structure to the overall test. Swain (1982) and colleagues provide an example of such a test, in which a variety of subthemes and language tasks are naturally linked to an overall theme, such as organizing a summer music festival or setting up a student job program. The soundness of this approach deserves more intensive examination since there would be a need to avoid item content bias that might be introduced.

(b) Four stages in test administration. In the Oral Interview of the U.S. Government, oral interaction skills are elicited in four stages, each of which has psychological, linguistic, and evaluative purposes. The warm-up is intended to put the examinee at ease and to familiarize him or her with the target language and with the interviewers. Next, the level check seeks to identify that proficiency level at which the examinee performs best (e.g., most comfortably and most satisfactorily). Then, at the probe stage, an effort is made to challenge the examinee with tasks thought to be just beyond this identified level, both to verify the examinee's maximum proficiency level and to demonstrate to the examinee what tasks he or she cannot yet perform. Finally, the wind-up presents tasks at the examinee's best performance level so that he or she finishes the test with a sense of accomplishment. (For further discussion and references on these stages in administration of both productive and receptive language tests, see Lowe [1981]; and Canale, forthcoming.)

(c) Adaptive testing procedures. In contrast to a traditional test, in which a group of examinees is required to respond to the same items regardless of the individuals' varying abilities to do so, an adaptive test is tailored during its actual administration according to each examinee's level of performance on successive items or groups of items. The Oral Interview procedure just described is one example of an adaptive testing procedure. Adaptive tests may also be delivered by computer. For example, an examinee who performed poorly on the first, say, five items of a reading test could be automatically administered five items different (and presumably easier) from those items automatically administered to another examinee who performed well on the first five. Automatic presentation of further groups of items would depend on each examinee's performance on the second group of five, and so on, until a consistent (and here different) level of performance had been confirmed for each student. Measurement experts such as Green (1983) and Wainer (1983) argue that such adaptive testing is not only feasible from measurement and technological viewpoints, but is highly desirable for a variety of reasons. One important reason is that an adaptive test allows the examinee to work at his or her own pace and provides repeated measurements at his or her performance level, and hence less room for measurement error. Another is that examinees are likely to find adaptive tests less boring, frustrating, and tiring than traditional ones, since most items should correspond to the individual examinee's performance level and since there are fewer items and less emphasis on speed.

(d) Criterion-referenced tests. Whereas a norm-referenced test consists of tasks designed to maximize differences in performance among examinees, a criterion-referenced test consists of whatever tasks (or sample of tasks) examinees must be able to perform for a given purpose and at a given proficiency level. In other words, a criterion-referenced test is designed to determine the extent to which a given examinee can or cannot perform a target (or criterion) task. Educational evaluation experts such as Popham (1975:134 ff.) often stress two main advantages of criterion-referenced tests over norm-referenced ones. First, the tasks on the former must represent, in a direct and theoretically sound manner, tasks that are crucial to performance of the criterion task; the tasks on a norm-referenced test need only produce

scores that show a high correlation with scores on the criterion task. Thus, while one could conceive of a norm-referenced test of oral interaction skills in which no examinee was ever required to speak, such a criterion-referenced test would be inconceivable. Second, the scores on a criterion-referenced test are intended to be directly interpretable in terms of some actual criterion performance; scores on a norm-referenced test are intended only to indicate how each examinee stands in comparison to a larger group of examinees, with little or no clear feedback as to the reasons for such relative standing. While criterion-referenced testing also has certain disadvantages (e.g., in requiring new statistical techniques for computing reliability estimates) and is not suitable or desirable in all testing situations (cf. Allen and Yen, 1979), it is nonetheless a worthwhile consideration in the testing of both productive and receptive language skills.

4. Implications for TOEFL Program Activities

This section suggests some activities addressing issues missed by this report that might be considered by the TOEFL program. The intent of these suggestions is not to provide firm prescriptions, but rather to mention possibile activities that are likely to have positive benefits for improving use of the existing TOEFL and for research and test development activities.

4.1 Study of the TOEFL Test Specifications and Continued Study of the Content Characteristics of TOEFL Items

The research described in this report did not refer to the TOEFL test specifications. While the investigators had access to the specifications and, indeed, reviewed them briefly in preparation for the research, there was no opportunity to investigate how the conclusions drawn from this report might be useful in reviewing the existing test specifications. Such a review would seem to be an important next step in applying the findings of this report. Attention should focus on the overlap and lack of overlap between the description of the TOEFL test given here and in the specifications; one valuable goal would be to identify ways in which the specifications might be revised to increase the naturalness of discourse occurring in the test.

Parallel with review of the specifications, further research and test development activities aimed at describing TOEFL items using communicative approaches should be considered. This effort is important in order to understand how the content characteristics of TOEFL items are related to examinees' performance on items.

4.2 Development Research on Thematic Presentation of Items and on New Item Types

The TOEFL program should consider ways for increasing the face validity and naturalness of language usage occurring in the TOEFL test. It is important to recognize that changes in the existing TOEFL cannot be undertaken in an abrupt fashion. Any changes that are considered need to be evaluated for their soundness with regard to several criteria, e.g., psychometric

integrity of the test and implications for program operational procedures, including test development procedures, costs, and TOEFL score report use.

The first step in this process is to study alternative ways in which the face and content validity of the existing TOEFL might be improved. It ought to prove useful to consider ways in which the contents of existing item types might be made more thematically relevant to college life and academic experience. For example, a series of revised statement items might have contents that are related to a few general themes appropriate to college experience, such as attending a lecture series, learning to use a library, and registering for course work.

It is important to note that a decision to use items with a shared thematic content may have disadvantages as well as advantages. Research would be needed to make sure that using items with shared thematic content would not lead to test content bias. By restricting the number of topics that items may refer to, it is possible that examinees' differential familiarity with these topics may have a systematic impact on differences in test scores.

A second step worth considering is the development of new item types that would feature more extended use of integrated proficiency skills. The results of the present review of the TOEFL test suggest that the breadth of language recognition competencies required of examinees is directly related to the length and textual complexity of TOEFL item stimuli. Accordingly, increasing the amount of language and the complexity of discourse in items is likely to benefit the range and nature of skills exercised in taking TOEFL. The process of suggesting new item types is nontrivial. The considerations raised by Bachman and Palmer (forthcoming) as factors influencing test performance will need to be considered. For example, increasing the amount of language in item stimuli and the number of questions, and hence distance between questions and stimuli, may have unexpected side effects on performance.

4.3 New Approaches to Assess Communicative Proficiency

The existing TOEFL, even with modifications, could not be expected to assess some important proficiency skills. For example, the test cannot test speaking ability or writing ability directly. Accordingly, efforts should be considered to map out additional competencies that may be tested by other instruments, such as the Test of Spoken English (TSE) and by prototype direct essay tasks currently being field tested. In addition, thought should be given to the desirability and feasibility of altogether new instruments to assess communicative skills not now tested.

4.4 Technology, Innovative Measurement and Adaptive Testing

Another consideration raised here concerns the importance of evaluating how advances in technology might aid the TOEFL program in developing new items and tests capable of assessing new ranges of communicative skills. The TOEFL program should take steps to enhance its access to information concerning new technological advances, involving microcomputers,

programming languages and software, video disks, voice synthesis, and other electronic technology, that could have an impact on development of new item types. It is important to note that some technologies, such as voice synthesis, are not yet feasible to implement easily on existing microcomputers because of memory limitations.

The interactive capabilities of microcomputers, plus the availability of video disks, may portend new possibilities for assessment of language proficiency skills not currently assessed by TOEFL. For example, it ought to prove feasible in technologically advanced countries to present examinees with video recordings of a person or persons conversing with each other, delivering a lecture, or asking questions of an examinee. One can envision an enhanced version of Section 1 of the current TOEFL that could be delivered by video disk. The use of a visual medium for test presentation has the advantage that it is standardized over examinees and would permit a richer range of questions to be asked about a discourse. A disadvantage might be that it could prove difficult to ensure the security of video-based items to the extent now possible with printed items.

Other possibilities involving the interactive capabilities of computers in language testing should be explored. For example, is it possible to design innovative conversation items that would permit assessment of examinees' ability to display sociolinguistic, discourse, and strategic competence? Obviously, in considering such suggestions, beyond construct and content validity issues, operational constraints in administering a test must be kept in mind.

Computer adaptive testing using conventional or enhanced TOEFL items is still another possibility. This area is currently under active investigation by the TOEFL program.

4.5 Validity Research

As this report demonstrates, it is possible to identify some ways in which the TOEFL examination resembles or differs from language use in authentic settings. A basic question that deserves research is how well the existing TOEFL, with its strengths and limitations, is capable of measuring language proficiency skills that are strongly related to performance in authentic language use situations. The skills required in taking TOEFL may or may not be predictive of a broader range of examinee language proficiency skills. This matter can only be resolved through empirical research on the question. An important step in pursuing such research will be to identify authentic, criterion language tasks and appropriate performance measures on these tasks. One valuable orientation toward the design of research studies of the sort that are needed can be found in the work of Bachman and Palmer (1981). These researchers have utilized a multitrait-multimethod approach to investigate the convergent and discriminant validity of batteries of proficiency tests designed to measure a variety of communicative proficiency skills.

4.6 Staying Abreast with the Assessment Field

In closing this report, it is important to endorse the significant benefits that language proficiency test developers can derive from contact with new research in the language proficiency assessment field. The field is undergoing an evolution as a result of attempts to increase the breadth of language skills that instruments might test. It is to the obvious advantage of test producers that they stay informed of developments. Contact with the field and evaluation of its advances with regard to the validity and appropriate use of existing tests is not unidirectional. As this report illustrates, there is merit in evaluating a proficiency test in light of communicative approaches to language assessment. The fact that a communicative approach leads to identification of many important skills not reflected on a test does not imply that simple remedies are at hand to improve existing tests. Clearly, it is of fundamental importance to test use that improvements in testing be identified, but it is far easier to identify skills not assessed by a test than it is to devise practical and scientifically sound ways in which to legitimately assess skills that are not currently assessed. We should always expect that language assessment theory will be ahead of language assessment practice. Amidst all the leading theoretical developments in proficiency assessment, some issues and approaches to assessment will be more tractable than others. It is important that proficiency test developers contribute to the cutting edge of these concerns, even though there can be no guarantees that new theory and assessment approaches will necessarily always lead to improvement in testing practices. Many steps are involved in striving for the goal of improved tests. This report illustrates ways in which theory might help in evaluating and improving existing instruments and testing practices.

The ways in which testing programs, such as the TOEFL program, respond to advances in language proficiency assessment theory will be of significance to language assessment research. The responses will inform researchers as to the gap that remains between theory and practice, and to the problems encountered in enacting testing practices that are responsive to advances in theory. This feedback may, in turn, contribute to the design of more adequate models for language proficiency assessment.

References

Allen, M. J., & Yen, W. M. (1979). Introduction to measurement theory. Monterey, CA: Brooks-Cole Publishing.

Austin, J. L. (1962). How to do things with words. Cambridge, MA: Harvard University Press.

Bachman L., & Palmer A. (1981). A multitrait-multimethod investigation into the construct validity of six tests of speaking and reading. In A. Palmer, P. Groot, & G. Trosper (Eds.), The construct validation of tests of communicative competence. Washington, DC: Teachers of English to Speakers of Other Languages.

Bachman, L., & Palmer, A. (forthcoming). Basic concerns in language test validation. Reading, MA: Addison-Wesley.

Bolinger, D. & Sears, D. (1981). Aspects of language. New York: Harcourt, Brace, Jovanovich, Inc..

Britton (1970)

Bruner, J. S. (1975). Language as an instrument of thought. In A. Davies (Ed.), Problems of language and learning. London: Heinemann.

Buros, O. K. (1977). Fifty years in testing: Some reminiscences, criticisms, and suggestions. Educational Research, 6, 9-15.

Burt, M. K., Dulay, H. C., & Hernandez-Chavez, E. (1976). Bilingual syntax measure technical manual. New York: Psychological Corporation.

Canale, M. (1984). Testing in a communicative approach. In G. A. Jarvis (Ed.), The challenge for excellence in foreign language education. Middlebury, VT: The Northeast Conference.

Canale, M. (1983a). On some dimensions of language proficiency. In J. W. Oller, Jr. (Ed.), Issues in language testing research. Rowley, MA: Newbury House Publishers.

Canale, M. (1983b). From communicative competence to communicative language pedagogy. In J. C. Richards and R. Schmidt (Eds.), Language and communication. London: Longman.

Canale, M. (Forthcoming). Considerations in the testing of reading listening proficiency. Foreign Language Annals (fall 1984).

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1, 1-47.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Clark, J. (1980). Toward a common measure of speaking proficiency. In J. Frith (Ed.), Measuring spoken language proficiency. Washington, DC: Georgetown University Press.

Clark, J. L. D. (1972). Foreign language testing: theory and practice. Philadelphia, PA: The Center for Curriculum Development.

Clark, J. L. D. (Ed.). (1978). Direct testing of speaking proficiency: Theory and application. Princeton, NJ: Educational Testing Service.

Clark, J. L. D., & Swinton, S. S. (1980). The Test of Spoken English as a measure of communicative ability in English-medium instructional settings. Princeton, NJ: Educational Testing Service. (TOEFL Research Report 7).

Clark, H., & Clark, E. (1977). Psychology and language. New York: Harcourt, Brace, Jovanovich, Inc..

Cummins, J. (1983). Language proficiency and academic achievement. In J. W. Oller, Jr. (Ed.), Issues in language testing research. Rowley, MA: Newbury House Publishers.

Cummins, J. (1984). Wanted: A theoretical framework for relating language proficiency to academic achievement among bilingual students. In C. Rivera (Ed.), Language proficiency and academic achievement. Clevedon, England: Multilingual Matters.

Deyhle, D. (1983). Between games and failure: A micro-ethnographic study of Navajo students and testing. Curriculum Inquiry, 13(4).

Duran, R. P. (1984). Implications of communicative competence research for integrative proficiency testing. In C. Rivera (Ed.), Communicative competence approaches to language proficiency assessment: Research and application. Clevedon, England: Multilingual Matters.

Educational Testing Service. (1980). Understanding TOEFL: Test Kit 1. Test of English as a Foreign Language Program. Princeton, NJ.

Educational Testing Service. (1983). Test and Score Manual, 1983 Edition, Test of English as a Foreign Language Program. Princeton, NJ.

Fodor, J., Bever, T., & Garrett, M. (1974). The psychology of language. New York: McGraw-Hill.

Francis, W. N. (1958). The structure of American English. New York: The Ronald Press Company.

Garvin, P. (1978). An empirilist epistemology for linguistics. In M. Paradis (Ed.), The Fourth LACUS Forum 1977. Columbia, SC: Hornbeam Press.

Green, B. (1983). The promise of tailored tests. In H. Wainer & S. Messick (Eds.), Principals of modern psychological measurement. (A Festschrift for Frederic M. Lord.) Hillsdale, NJ: Lawrence Erlbaum Associates.

Gumperz, J. J. (1982). Discourse strategies. Cambridge, MA; Cambridge University Press.

Gumperz, J. J. (1971). Language in social groups. Essays by John J. Gumperz. Selected and introduced by Anwar S. D. L.. Stanford, CA: Stanford University Press.

Gumperz, J. J., & Kaltman, H. (1980). Prosody, linguistic diffusion and conversational inference. In Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society. Berkeley, CA: Berkeley Linguistics Society, 44-65.

Hale, G., & Hinofotis, F. (1981). New directions in English language testing. Princeton, NJ: Test of English as a Foreign Language Program. Unpublished manuscript.

Halliday, M. A. K. (1973). Explorations in the functions of language. New York: Elsevier North-Holland, Inc.

Hatch, E. M. (1983). Psycholinguistics. A second language perspective. Rowley, MA: Newbury House Publishers, Inc.

Hudson, R. A. (1980). Sociolinguistics. Cambridge, MA: Cambridge University Press.

Hymes, D. (1974). Foundations in sociolinguistics. An ethnogrophic approach. Philadelphia, PA: University of Pennsylvania Press.

Jones, R. (1983, March). Current research in second language performance testing. Keynote address presented at the Conference on Second Language Performance Testing, University of Ottawa and Carleton University, Ottawa, Ont..

Joos, M. (1967). The five clocks. New York: Harcourt, Brace, Jovanovich, Inc..

Kaplan, R. (1966). Cultural thought patterns in intercultural education. Language Learning, 16, 1-20.

Krashen, S. (1982). Principles and practice in second language acquisition. Oxford: Pergamon.

Leech, G., & Svartvik, J. (1975). A communicative grammar of English. Singapore: Longman.

Lowe, P. (1981). Structure of the oral interview and content validity. In A. S. Palmer, P. J. M. Groot, & G. A. Trosper (Eds.), The construct validation of tests of communicative competence. Washington, DC: TESOL.

Lowe, P. (forthcoming). "The" question: A response to Michael Canale. Foreign Language Annals, (fall 1974).

Malinowski, B. (1944). A scientific thory of culture and other essays. Chapel Hill: The University of North Carolina Press.

Mathiot, M. (1970). The semantic and cognitive domains of language. In P. Garvin (Ed.), Cognition: A multiple view. New York: Spartan Books.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. Language Learning, 33(2), 135-158.

Messick, S. (1981). Evidence and ethics in the evaluation of tests. Educational Researcher, 10(9), 9-20.

Munby, J. (1978) Communicative syllabus design. Cambridge, MA; Cambridge University Press.

Oller, J. W., & Spolsky, B. (1979). The Test of English as a Foreign Language (TOEFL). In B. Spolsky (Ed.), Advances in language testing. Volume 1. Washington, DC: Center for Applied Linguistics.

Penfield, J. (1979). Essay writing for the African world. Unpublished manuscript.

Popham, W. J. (1975). Educational evaluation. Englewood Cliffs, NJ: Prentice-Hall.

Savignon, S. (1983) Communicative competence: Theory and classroom practice. Reading, MA: Addison Wesley.

Scott, M., & Madsen, H. (1983). The influence of retesting on test affect. In J. Oller (Ed.), Issues in language testing research. Rowley, MA: Newbury House Publishers.

Shoemaker, D. (1980). Improving achievement testing. Educational Policy Analysis, 2(6), 37-49.

Spolsky, B. (1979). Linguistics and language testers. In B. Spolsky (Ed.), Advances in language testing, Volume 1. Washington, DC: Center for Applied Linguistics.

Swain, M. (1982, December). Large scale communnicative language testing: A case study. Plenary paper presented at the International Symposium on Language Testing. Hong Kong: University of Hong Kong, Mimeo.

Vygotsky, L. S. (1962). Thought and language. Cambridge, MA: MIT Press.

Wainer, H. (1983). On item response theory and computerized adaptive tests. The Journal of College Admissions, 28(4), 9-16.

Wesche, M. B. (1981). Communicative testing in a second language. The Canadian Modern Language Review, 37(3), 551-571.

Yorio, C. A. (1980). Conventionalized language forms and the development of communicative competence. TESOL Quarterly, 14(4), 433-442.

APPENDIX A

Sample TOEFL Test

TEST OF ENGLISH AS A FOREIGN LANGUAGE

General Directions

This is a test of your ability to use the English language. It is divided into three sections, some of which have more than one part. Each section or part of the test begins with a set of specific directions that include sample questions. Be sure you understand what you are to do before you begin to work on a section.

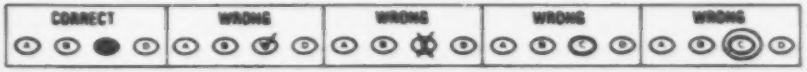
The supervisor will tell you when to start each section and when to go on to the the next section. You should work quickly but carefully. Do not spend too much time on any one question. If you finish a section early, you may review your answers on that section only. You may not go on to the next section and you may not go back to a section you have already worked on.

You will find that some of the questions are more difficult than others, but you should try to answer every one. Your score will be based on the number of correct answers you give. If you are not sure of the correct answer to a question, make the best guess that you can. It is to your advantage to answer every question, even if you have to guess the answer.

Do not mark your answers in this test book. You must mark all of your answers on the separate answer sheet that is inside this test book. When you mark your answer to a question on your answer sheet, you must:

- Use the pencil you have been given or another medium-soft (#2 or HB) black lead pencil.
- Carefully and completely blacken the oval corresponding to the answer you choose for each question. Be sure to mark your answer in the row with the same number as the number of the question you are answering.
- Mark only one answer to each question.
- Completely fill the oval with a heavy, dark mark so that you cannot see the letter inside the oval. Light or partial marks may not be read properly by the scoring machine.
- Erase all extra marks completely and thoroughly. If you change your mind about an answer after you have marked it on your answer sheet, completely erase your old answer and then mark your new answer.

The examples below show you the correct and wrong ways of marking an answer sheet. Be sure to fill in the ovals on your answer sheet the correct way.



Some or all of the passages for this test have been adapted from published material to provide the examinee with significant problems for analysis and evaluation. To make the passages suitable for testing purposes, the style, content, or point of view of the original may have been altered in sbme cases. The ideas contained in the passages do not necessarily represent the opinions of the TOEFL Policy Council or Educational Testing Service.

679877
VV23P5,6 pp.25 qtn.150
Printed in U.S.A.

BEST COPY AVAILABLE

Print your full name here _____
(last) (first) (middle)

FORM 3FATF5

00?R1

TEST OF ENGLISH AS A FOREIGN LANGUAGE

Read the directions on the back cover.

Do not break the seals until you are told to do so.

This test book was used at the May 14, 1983, TOEFL administration. It is distributed by the TOEFL program office only to examinees who took the test on May 14, 1983.

UNAUTHORIZED REPRODUCTION IS PROHIBITED.

Copyright © 1983 by Educational Testing Service. All rights reserved.
Princeton, N.J. 08541

SECTION I

LISTENING COMPREHENSION

In this section of the test, you will have an opportunity to demonstrate your ability to understand spoken English. There are three parts to this section, with special directions for each part.

Part A

Directions: For each question in Part A, you will hear a short statement. The statements will be spoken just one time. They will not be written out for you, and you must listen carefully to understand what the speaker says.

After you hear a statement, read the four sentences in your test book, marked (A), (B), (C), and (D), and decide which one is closest in meaning to the statement you heard. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen so that the letter inside the oval cannot be seen.

Example I — Sample Answer

You will hear:

You will read:

(A) John does better in his studies than James.
(B) James is bigger than his brother John.
(C) John has only one brother.
(D) The teacher likes James better than John.

Sentence (A), "John does better in his studies than James," means most nearly the same as the statement "John is a better student than his brother James." Therefore, you should choose answer (A).

Example II — Sample Answer

You will hear:

You will read:

(A) The traffic isn't bad today.
(B) The trucks weigh a lot.
(C) There are a lot of trucks on the highway.
(D) The highway has been closed to heavy trucks.

Sentence (C), "There are a lot of trucks on the highway," is closest in meaning to the sentence "The truck traffic on this highway is so heavy I can barely see where I'm going." Therefore, you should choose answer (C).

1. (A) Go directly to the post office when class is over.
   (B) Let's first straighten up the classroom and then go to the post office.
   (C) That's the most direct way to the post office from our class.
   (D) The post office is straight ahead of the classroom building.

2. (A) I don't think that algebra is hard.
   (B) I like algebra better than geometry.
   (C) Geometry isn't difficult for me.
   (D) Geometry is easier for me than algebra.

3. (A) Greg believed he could do it alone.
   (B) Greg thought he'd cut himself.
   (C) Greg thought he was selfish.
   (D) Greg alone believed it could be done.

4. (A) After it rained, he washed his car.
   (B) He was unable to wash his car because it was raining.
   (C) It began to rain right after he washed his car.
   (D) He had to finish washing his car in the rain.

5. (A) Don't make noise in the kitchen.
   (B) You may not cook here.
   (C) They were quiet when they ate.
   (D) These are homemade cookies.

6. (A) You should call Margaret soon.
   (B) Margaret will be better later on.
   (C) It's too late to call on Margaret now.
   (D) Margaret is the best person to tell.

7. (A) He never walks to the library at night.
   (B) There is only one librarian here at night.
   (C) The library is the only place to study.
   (D) He never works in the library in the daytime.

8. (A) How was your dinner?
   (B) Please have dinner with us.
   (C) We had dinner together.
   (D) Will there be four of us for dinner?

9. (A) Jerry dislikes the clothes he has.
   (B) Jerry doesn't like doing his laundry.
   (C) Jerry hates to take showers.
   (D) Jerry's clothes don't need ironing.

10. (A) Debbie checked with her son's doctor.
    (B) Debbie sent her son for a checkup.
    (C) Debbie paid her son's doctor.
    (D) Debbie wrote a note to the doctor's son.

11. (A) The pool was scheduled to open on Tuesday.
    (B) The pool is opening today.
    (C) The pool will open tomorrow.
    (D) The pool should be open on Saturday.

12. (A) Nelson Studios took the picture for my passport.
    (B) I studied the photograph of the port.
    (C) I took my passport to the studios.
    (D) I pass by Nelson Studios on my way to work.

13. (A) I told you to see a lot of museums.
    (B) You've taught me a great deal here.
    (C) People say that you know this place well.
    (D) Many museums are like this one, you know.

14. (A) I'd expected Linda to give a performance.
    (B) Linda hadn't been expecting to perform.
    (C) I'd expected Linda to do things differently.
    (D) Linda had expected me to be there.

15. (A) You'll probably finish in time to help Dorothy.
(B) Dorothy would be a great help to you.
(C) Dorothy won't finish without your help.
(D) There's no time for Dorothy to help you.

16. (A) I have no supplies.
(B) I just left the supply store.
(C) I just found a supply.
(D) I went out to get supplies.

17. (A) You don't think the seminar is fascinating, do you?
(B) The seminar will continue while we are eating.
(C) I find the seminar extremely interesting.
(D) The dissemination of information is fast, isn't it?

18. (A) She didn't use the record player afterwards.
(B) She didn't play all her old records.
(C) She bought the same kind of record player.
(D) She kept her old record player.

19. (A) The encyclopedias were on order.
(B) The encyclopedias weren't checked out.
(C) The encyclopedias weren't any good.
(D) The encyclopedias were improperly arranged.

20. (A) We broke two cups.
(B) We left at two o'clock.
(C) We divided into two groups.
(D) We met two people.

Part B

Directions: In Part B you will hear short conversations between two speakers. At the end of each conversation, a third voice will ask a question about what was said. The question will be spoken just one time. After you hear a conversation and the question about it, read the four possible answers in your test book and decide which one is the best answer to the question you heard. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen.

Example I

Sample Answer

You will hear:

You will read:
(A) Read a book.
(B) Write a composition.
(C) Talk about a problem.
(D) Listen to the radio.

From the conversation you know that the assignment is to listen to a radio program and be ready to talk about it. The best answer, then, is (D), "Listen to the radio." Therefore, you should choose answer (D).

21. (A) In a department store.
(B) In a bank.
(C) At a tourist bureau.
(D) At a hotel.

22. (A) Eat before seeing the movie.
(B) See the movie immediately.
(C) Get the first theater seat.
(D) Stay in town for a while.

23. (A) The winter has just begun.
(B) Once it starts, it'll snow a lot.
(C) They're ready for the snow.
(D) It has been snowing for some time.

24. (A) Traveling a lot.
(B) Getting a lot of exercise.
(C) Working too hard.
(D) Waiting for the train.

25. (A) He can't find his new building.
(B) He had a bigger apartment before.
(C) He's not accustomed to the large building.
(D) He's having a hard time finding an apartment.

26. (A) Find the trouble.
(B) Carry the parts outside.
(C) Practice working together.
(D) Follow the directions.

27. (A) He fixes bicycles.
(B) He raises sheep.
(C) He sells chairs.
(D) He's a gardener.

28. (A) It doesn't matter which color the man chooses.
(B) It's a difficult decision.
(C) She doesn't like either color.
(D) The man should choose a different room.

29. (A) She'd like the store to send it to her.
(B) It will arrive next week.
(C) It must be wrapped quickly.
(D) She'll take it with her to save trouble.

30. (A) They want to go downtown.
(B) He wants to go to the park, but she doesn't.
(C) He doesn't know where to park the car.
(D) He wants to find out the location of the park.

31. (A) Try a new ribbon.
(B) Help her type the paper.
(C) Get another typewriter.
(D) Change the paper.

32. (A) Tom is unable to hear well.
(B) Tom didn't say anything at the meeting.
(C) Tom doesn't listen to him.
(D) Tom went out before the meeting was over.

33. (A) Help fill up the boxes.
(B) Take some of the boxes.
(C) Look for something else.
(D) Make sure her hands are clean.

34. (A) He wants the others to follow him.
(B) He must study the animals he caught.
(C) He will catch up with them later.
(D) He is behind in his schoolwork.

35. (A) You should believe everything you read.
(B) She thinks the book is excellent.
(C) She wonders which newspaper he reads.
(D) Reaction to the book has been varied.

BEST COPY AVAILABLE

Part C

Directions: In this part of the test, you will hear several short talks and conversations. After each talk or conversation, you will be asked some questions. The talks and questions will be spoken just one time. They will not be written out for you, so you will have to listen carefully to understand what the speaker says.

After you hear a question, read the four possible answers in your test book and decide which one is the best answer to the question you heard. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen.

Listen to this sample talk.

You will hear:

Now look at the following example.

You will hear:

You will read:
(A) By plane.
(B) By ship.
(C) By train.
(D) By bus.

Sample Answer

Ⓐ ● Ⓒ Ⓓ

The best answer to the question "How did people generally arrive at Ellis Island?" is (B), "By ship." Therefore, you should choose answer (B).

Now look at the next example.

You will hear:

You will read:
(A) New immigrants.
(B) International traders.
(C) Fishermen.
(D) Tourists.

Sample Answer

Ⓐ Ⓑ Ⓒ ●

The best answer to the question "Who visits Ellis Island today?" is (D), "Tourists." Therefore, you should choose answer (D).

36. (A) At a telephone laboratory.
(B) At the library.
(C) On Martha's Vineyard.
(D) In a lecture hall.

37. (A) It was settled more than 300 years ago.
(B) Alexander Graham Bell visited there.
(C) A large number of its residents were deaf.
(D) Each family living there had many children.

38. (A) They inherited deafness.
(B) An epidemic struck the island.
(C) The climate caused hearing loss.
(D) It was an unlucky place.

39. (A) Two.
(B) Seventeen.
(C) Twenty-five.
(D) Forty.

40. (A) Establish his laboratory.
(B) Have a vacation.
(C) Study deafness among the families.
(D) Visit members of his family.

41. (A) The patterns of marriage have changed.
(B) Many deaf people have regained their hearing.
(C) Most of the original population has left the island.
(D) The island has become famous for its research facilities.

42. (A) San Francisco.
(B) Forest fires.
(C) Redwood trees.
(D) Survival skills.

43. (A) In Muir Woods.
(B) Near Los Angeles.
(C) In San Francisco.
(D) Along the northern California coast.

44. (A) It has no admission fee.
(B) It is near San Francisco.
(C) It has a good view of the coast.
(D) It can be seen in one hour.

45. (A) 350 years.
(B) 400 years.
(C) 800 years.
(D) 2,000 years.

46. (A) Absence of natural enemies.
(B) Resistant bark and damp climate.
(C) Coastal isolation.
(D) Cool weather and daily fog.

47. (A) Book cataloguers.
(B) People shelving books.
(C) People reading magazines.
(D) Students doing research.

48. (A) Very shortly.
(B) After everyone has finished.
(C) Tomorrow night.
(D) In a few days.

49. (A) Close their test books now.
(B) Return the next day to finish.
(C) Put books back where they belong.
(D) Check to see if they have their books.

50. (A) They can be taken out overnight.
(B) They will be held overnight.
(C) They need to be returned now.
(D) They are on a special shelf.

[redacted]

THE NEXT PART OF THE TEST IS SECTION 2. TURN TO THE DIRECTIONS FOR SECTION 2 IN YOUR TEST BOOK, READ THEM, AND BEGIN WORK.
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.

STOP STOP STOP STOP STOP STOP STOP

SECTION 2

STRUCTURE AND WRITTEN EXPRESSION

Time—25 minutes

This section is designed to measure your ability to recognize language that is appropriate for standard written English. There are two types of questions in this section, with special directions for each type.

Directions: Questions 1-15 are incomplete sentences. Four words or phrases, marked (A), (B), (C), and (D), are given beneath each sentence. You are to choose the one word or phrase that best completes the sentence. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen so that the letter inside the oval cannot be seen.

Example I

Mt. Hood —— in the state of Oregon.

(A) although
(B) and
(C) is
(D) which

Sample Answer

Ⓐ Ⓑ ● Ⓓ

In English, the sentence should read, "Mt. Hood is in the state of Oregon." Therefore, you should choose (C).

Example II

—— most important event in San Francisco's history was the disastrous earthquake and fire of 1906.

(A) The
(B) It was the
(C) That the
(D) There was a

Sample Answer

● Ⓑ Ⓒ Ⓓ

In English, the sentence should read, "The most important event in San Francisco's history was the disastrous earthquake and fire of 1906." Therefore, you should choose (A).

As soon as you understand the directions, begin work on the questions.

1. Conifers first appeared on the Earth —— the early Permian period, some 270 million years ago.
   (A) when
   (B) or
   (C) and
   (D) during

2. There are very few areas in the world —— be grown successfully.
   (A) where apricots can
   (B) apricots can
   (C) apricots that can
   (D) where can apricots

3. —— a baby turtle is hatched, it must be able to fend for itself.
   (A) Not sooner than
   (B) No sooner
   (C) So soon that
   (D) As soon as

4. Tungsten, a gray metal with the ——, is used to form the wires in electric light bulbs.
   (A) point at which it melts is the highest of any metal
   (B) melting point is the highest of any metal
   (C) highest melting point of any metal
   (D) metal's highest melting point of any

5. Rattan comes from —— of different kinds of palms.
   (A) its reedy stems
   (B) the reedy stems
   (C) the stems are reedy
   (D) stems that are reedy

6. At thirteen —— at a district school near her home, and when she was fifteen, she saw her first article in print.
   (A) the first teaching position that Mary Jane Hawes had
   (B) the teaching position was Mary Jane Hawes' first
   (C) when Mary Jane Hawes had her first teaching position
   (D) Mary Jane Hawes had her first teaching position

7. Vitamin C, discovered in 1932, —— first vitamin for which the molecular structure was established.
   (A) the
   (B) was the
   (C) as the
   (D) being the

8. The behavior of gases is explained by —— the kinetic theory.
   (A) what scientists call
   (B) what do scientists call
   (C) scientists they call
   (D) scientists call it

9. Ironically, sails were the salvation of many steamships —— mechanical failures.
   (A) they suffered
   (B) suffered
   (C) were suffered
   (D) that had suffered

10. —— some mammals came to live in the sea is not known.
    (A) Which
    (B) Since
    (C) Although
    (D) How

11. —— their nests well, but also build them well.
    (A) Not only brown thrashers protect
    (B) Protect not only brown thrashers
    (C) Brown thrashers not only protect
    (D) Not only protect brown thrashers

12. The name Nebraska comes from the Oto Indian word "nebrathka," —— flat water.
    (A) to mean
    (B) meaning
    (C) it means
    (D) by meaning

13. Biochemists use fireflies to study bioluminescence, ——.
    (A) the heatless light given off by certain plants and animals
    (B) certain plants and animals give off the heatless light
    (C) which certain plants and animals give off the heatless light
    (D) is the heatless light given off by certain plants and animals

14. Rich tobacco and champion race horses have —— of Kentucky.
    (A) long been symbols
    (B) been long symbols
    (C) symbols been long
    (D) long symbols been

15. Today's libraries differ greatly from ——.
    (A) the past
    (B) those of the past
    (C) that are past
    (D) those past

Directions: In questions 16-40 each sentence has four words or phrases underlined. The four underlined parts of the sentence are marked (A), (B), (C), and (D). You are to identify the one underlined word or phrase that should be corrected or rewritten. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen.

Example I — Sample Answer ● Ⓑ Ⓒ Ⓓ

Much (A) federal and industrial experts are certain (B) that meat shortages (C) will cause an enormous increase (D) in the consumption of fish and other sea products.

Answer (A), the underlined word much, would not be accepted in carefully written English; the word many is used with the plural experts. Therefore, the sentence should read, "Many federal and industrial experts are certain that meat shortages will cause an enormous increase in the consumption of fish and other sea products." To answer the question correctly, you would choose (A).

Example II — Sample Answer Ⓐ Ⓑ ● Ⓓ

It was during the 1920's (A) that the friendship (B) between Hemingway and Fitzgerald reached their (C) highest (D) point.

Answer (C), the underlined word their, would not be accepted in carefully written English; the singular form its should be used because friendship is singular. Therefore, the sentence should read, "It was during the 1920's that the friendship between Hemingway and Fitzgerald reached its highest point." To answer the question correctly, you should choose (C).

As soon as you understand the directions, begin work on the questions.

16. Edna Ferber told (A) the story (B) of her life (C) in two book (D).

17. The period (A) of a quarantine depends (B) to the amount of time necessary (C) for protection against (D) the spread of a particular disease.

18. By (A) 1642 all towns (B) in the colony of Massachusetts was (C) required by law to have (D) schools.

BEST COPY AVAILABLE

BEST COPY AVAILABLE

19. The bobwhite is the kind only of quail native to the area east of the Mississippi River.
    A B C D

20. June bugs they often cause damage by stripping the young leaves from trees and shrubs.
    A B C D

21. Since beginning of photography, inventors have tried to make photographs that duplicate natural colors.
    A B C D

22. Artificial rubies and sapphires have the same hard and composition as the real stones.
    A B C D

23. Although they sleep most of the winter, chipmunks are very active in summer when they gather and carry food in their cheek pouches and storing it underground.
    A B C D

24. Manufacturers use both chemical or mechanical methods to obtain clear glue.
    A B C D

25. The eel larva looks alike a thin willow leaf and is as transparent as glass.
    A B C D

26. The sun has always been a important guide to direction.
    A B C D

27. The manner in which fuel enters a diesel engine is the primary factor that affects its efficiently.
    A B C D

28. Traces of radon are found in the air in various amounts, according the weather.
    A B C D

29. Basic knowledge of mathematic and electronics was used to develop the high-speed electronic computer.
    A B C D

30. The field of guidance and counseling was still in its infancy in 1914 when Orie Hatcher entered them.
    A B C D

31. There are vineyards in California that product some of the best wine in the world.
    A B C D

32. The formation of snow must be occurring slowly, in calm air, and at a temperature near the freezing point.
    A B C D

33. Thomas Jefferson skillfully organized his supporter in Congress into a strong political group.
    A B C D

34. The greenest and plentifulest leaves are the leaves of grasses.
    A B C D

35. W. H. Auden's subtle use of everyday "nonpoetic" language and conversational rhythms great influenced modern poetry.
    A B C D

36. Long before the dawn of recorded history, humans celebrated to harvest their crops.
    A B C D

37. In 1931, Duke Ellington broke the traditional three-minutes time limit set for commercial records.
    A B C D

38. The basilisk lizard can run on the hind legs at speeds up to seven miles per hour.
    A B C D

39. Dairy farm is carried on in all parts of the United States.
    A B C D

40. Duckbill platypuses eat to half their weight each day, and at times even more.
    A B C D

THIS IS THE END OF SECTION 2

IF YOU FINISH BEFORE TIME IS CALLED, CHECK YOUR WORK ON SECTION 2 ONLY.
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.
THE SUPERVISOR WILL TELL YOU WHEN TO BEGIN WORK ON SECTION 3.

3 • 3 • 3 • 3 • 3 • 3 • 3

SECTION 3

READING COMPREHENSION AND VOCABULARY

Time—45 minutes

This section is designed to measure your ability to understand various kinds of reading materials, as well as your ability to understand the meaning and use of words. There are two types of questions in this section, with special directions for each type.

Directions: In questions 1-30 each sentence has a word or phrase underlined. Below each sentence are four other words or phrases, marked (A), (B), (C), and (D). You are to choose the one word or phrase that best keeps the meaning of the original sentence if it is substituted for the underlined word or phrase. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter you have chosen so that the letter inside the oval cannot be seen.

Example

The ordinary land snail moves at the rate of about two inches per minute.

(A) expert
(B) active
(C) common
(D) colorful

Sample Answer

Ⓐ Ⓑ ● Ⓓ

The best answer is (C) because "The common land snail moves at the rate of about two inches per minute" is closest in meaning to the original sentence, "The ordinary land snail moves at the rate of about two inches per minute." Therefore, you should choose answer (C).

As soon as you understand the directions, begin work on the questions.

1. In masculine rhyme, the end sounds of stressed syllables are repeated.
   (A) dominant
   (B) vowel
   (C) hard
   (D) final

2. About a quarter of the workers in the United States are employed in factories.
   (A) third
   (B) fourth
   (C) tenth
   (D) fifteenth

3. Bats are extremely shy creatures and avoid humans if at all possible.
   (A) timid
   (B) clean
   (C) private
   (D) noisy

4. Many kinds of seaweed grow along the Rhode Island seashore.
   (A) bank
   (B) coast
   (C) canal
   (D) gulf

5. Academic records from other institutions often become part of a university's official file and can neither be returned to a student nor duplicated.
   (A) borrowed
   (B) purchased
   (C) copied
   (D) rewritten

GO ON TO THE NEXT PAGE

BEST COPY AVAILABLE

3 • 3 • 3 • 3 • 3 • 3 • 3

6. Ammonia, one of the earliest known nitrogen compounds, was originally produced by distilling organic materials.
   (A) masses
   (B) fabrics
   (C) substances
   (D) liquids

7. Loud noises can be annoying.
   (A) hateful
   (B) painful
   (C) unnerving
   (D) irritating

8. The sensation of a "lump in one's throat" arises from an increased flow of blood into the tissues of the pharynx and larynx.
   (A) explanation
   (B) disease
   (C) feeling
   (D) unpleasantness

9. The apparently homogeneous Dakota grasslands are actually a botanical garden of more than 400 types of grasses.
   (A) seemingly
   (B) comparatively
   (C) dazzlingly
   (D) strangely

10. Photojournalist Margaret Bourke-White became famous for her coverage of significant events during the Second World War.
    (A) usage
    (B) camouflage
    (C) collage
    (D) reportage

11. Almost all economists agree that nations gain by trading with one another.
    (A) cooperate
    (B) profit
    (C) become more stable
    (D) become more dependent

12. The Cheyenne Indians were considered spectacular riders and fierce warriors.
    (A) hunters
    (B) defenders
    (C) enemies
    (D) fighters

13. It has been suggested that people who watch television incessantly may become overly passive.
    (A) seriously
    (B) skeptically
    (C) constantly
    (D) arbitrarily

14. While serving in the Senate in the early 1970's, Barbara Jordan supported legislation to ban discrimination and to deal with environmental problems.
    (A) list
    (B) forbid
    (C) handle
    (D) investigate

15. Aneroid barometers are able to show much slighter changes in the atmosphere than mercury barometers.
    (A) smaller
    (B) thinner
    (C) more sudden
    (D) more narrow

16. In a bullfight, it is the movement, not the color, of objects that arouses the bull.
    (A) confuses
    (B) excites
    (C) scares
    (D) diverts

GO ON TO THE NEXT PAGE

3 • 3 • 3 • 3 • 3 • 3 • 3

17. Sleep is associated with characteristic electrical rhythms in the brain.
    (A) cells
    (B) cues
    (C) shocks
    (D) patterns

18. Philip Roth was hailed as a major new author in 1960.
    (A) published
    (B) challenged
    (C) acclaimed
    (D) guided

19. Below 600 feet, ocean waters range from dimly lit to completely dark.
    (A) perversely
    (B) erratically
    (C) faintly
    (D) sufficiently

20. During their heyday, showboats were popular and generally prosperous.
    (A) golden age
    (B) infancy
    (C) summer voyages
    (D) revivals

21. Light rays are turned aside by the intense gravitational field surrounding a black hole.
    (A) heightened
    (B) deflected
    (C) rotated
    (D) created

22. The gar is a fish with a long, slender body and scales as hard as flint.
    (A) flat
    (B) straight
    (C) slim
    (D) fragile

23. Francis Scott Key wrote the words to "The Star-spangled Banner" after witnessing the unsuccessful attack on Fort McHenry.
    (A) participating in
    (B) observing
    (C) hearing about
    (D) resisting

24. Human facial expressions differ from those of animals in the degree to which they can be deliberately controlled and modified.
    (A) both
    (B) noticeably
    (C) intentionally
    (D) absolutely

25. If wool is submerged in hot water, it tends to shrink.
    (A) smell
    (B) fade
    (C) unravel
    (D) contract

26. The Constitution's vague nature has given it the flexibility to be adapted when circumstances change.
    (A) imprecise
    (B) diffuse
    (C) unpolished
    (D) elementary

100

COBA YAV

3 • 3 • 3 • 3 • 3 • 3 • 3

27. In the eighteenth century the heading of "natural philosophy" encompassed all of the sciences.
    (A) category
    (B) teachings
    (C) ideology
    (D) leaders

28. By today's standards, early farmers were imprudent because they planted the same crop repeatedly, exhausting the soil after a few harvests.
    (A) unwise
    (B) stubborn
    (C) tiresome
    (D) unscientific

29. The first step in planning a marketing strategy for a new product is to analyze the breakdown of sales figures for competitive products.
    (A) decrease in
    (B) reordering of
    (C) itemization of
    (D) collapse in

30. Georgia was colonized by a group of debtors from British prisons.
    (A) discovered
    (B) explored
    (C) settled
    (D) visited

BEST COPY AVAILABLE

101

3 • 3 • 3 • 3 • 3 • 3 • 3

Directions: The rest of this section is based on a variety of reading material (single sentences, paragraphs, and the like) followed by questions about the meaning of the material. For questions 31-60, you are to choose the one best answer, (A), (B), (C), or (D), to each question. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen.

Answer all questions following a passage on the basis of what is stated or implied in that passage.

Read the following passage.

> Despite all the atrocities falsely attributed to it, the gorilla is essentially a peace loving creature that would rather retreat than fight except when its life is threatened and retreat is impossible. In the wild it has never been seen eating meat, although some have learned to do so in captivity. Nor do gorillas seem to drink water in the wild; they apparently get what moisture they need from their diet of greenery and fruit.

Example I

Sample Answer

Gorillas have been known to eat meat only when they are

(A) in captivity
(B) in the wild
(C) engaged in fighting
(D) hiding from enemies

The passage says that some gorillas have learned to eat meat in captivity. Therefore, you should choose answer (A).

Example II

Sample Answer

Gorillas obtain most of the moisture they need from

(A) meat
(B) leaves and fruit
(C) small streams
(D) raids

The passage says that gorillas "apparently get what moisture they need from their diet of greenery and fruit." Therefore, you should choose (B) as the best completion of the sentence.

As soon as you understand the directions, begin work on the questions.

BEST COPY AVAILABLE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 31-34

Each variety of mosquito has its favored kind of water accumulation for breeding, and almost every imaginable type of still water has been used by at least one kind of mosquito to lay its eggs. After the eggs hatch, the larvae continue to be closely associated with the water's surface, hanging from the surface film and getting air through tubes that breaks the water's surface at the tail ends of their bodies. Because the larvae develop this way, they are never found in the open water of lakes where they would merely serve as fish food, or in places where they would be damaged by wave action or water currents.

31. According to the passage, what is true about the breeding habits of mosquitoes?
    (A) Different mosquitoes tend to have different kinds of breeding places.
    (B) Each mosquito usually breeds in several different places in one season.
    (C) A few mosquitoes constantly vary their breeding places.
    (D) Most mosquitoes mate in the same place in which they were bred.

32. According to the passage, most mosquito larvae develop
    (A) on plants near water
    (B) near sources of food
    (C) under waterproof sacs
    (D) in bodies of still water

33. According to the passage, most mosquito larvae breathe with
    (A) their wings
    (B) special tubes
    (C) their gills
    (D) modified mouths

34. According to the passage, mosquito larvae are never found in open water because they cannot
    (A) withstand much motion
    (B) find sufficient food there
    (C) obtain enough air there
    (D) tolerate too much moisture

BEST COPY AVAILABLE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 35-42

Decades before the American Revolution of 1776, Jesse Fish, a native New Yorker, retreated to an island off St. Augustine, Florida, to escape an unpleasant family situation. In time he became Florida's first orange baron and his oranges were in great demand in London throughout the 1770's. The English found them juicy and sweet and preferred them to other varieties, even though they had thin skins and were hard to peel.

There would probably have been other successful commercial growers before Fish if Florida had not been under Spanish rule for some two hundred years. Columbus first brought seeds for citrus trees to the New World and planted them in the Antilles. But it was most likely Ponce de León who introduced oranges to the North American continent when he discovered Florida in 1513. For a time, each Spanish sailor on a ship bound for America was required by law to carry one hundred seeds with him. Later, because seeds tended to dry out, all Spanish ships were required to carry young orange trees. The Spaniards planted citrus trees only for medicinal purposes, however. They saw no need to start commercial groves because oranges were so abundant in Spain.

35. What is the main topic of the passage?
    (A) The role of Florida in the American Revolution
    (B) The discovery of Florida by Ponce de León in 1513
    (C) The history of the cultivation of oranges in Florida
    (D) The popularity of Florida oranges in London in the 1770's

36. Jesse Fish came from
    (A) London
    (B) St. Augustine
    (C) the Antilles
    (D) New York

37. Jesse Fish went to Florida to
    (A) grow oranges commercially
    (B) buy an island off St. Augustine
    (C) get away from his family
    (D) work for the British government

38. Londoners liked the oranges grown by Jesse Fish because they
    (A) had a lot of juice
    (B) were not too sweet
    (C) were not hard to peel
    (D) had thin skins

39. Oranges were most probably introduced to Florida by
    (A) Jesse Fish
    (B) Ponce de León
    (C) Columbus
    (D) British sailors

40. According to the passage, Spanish vessels began to bring orange tree seedlings to North America when
    (A) the United States agricultural laws were revised
    (B) ambitious sailors began to smuggle seeds
    (C) doctors reported a lack of medical supplies
    (D) authorities realized that seeds did not travel well

41. According to the passage, Florida oranges were valued by the Spanish primarily
    (A) as a medium of exchange
    (B) for their unusual seeds
    (C) for their medical use
    (D) as a source of food for sailors

42. The Spaniards did not grow oranges commercially in the New World because
    (A) oranges tended to dry out during shipping
    (B) Florida oranges were very small
    (C) there was no great demand for oranges in Europe
    (D) oranges were plentiful in their home country

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 43-46

The very success of communications satellite systems has raised widespread concern about their future. Some countries are already using satellites for domestic communications in place of conventional telephone lines on land. Although this technique is extremely useful for linking widely scattered villages in remote or mountainous regions, in heavily built-up areas where extensive telephone and telegraph systems already exist domestic satellites (or "domsats") are seen by the land-line networks as unfair competition. Despite such opposition, domsats are gaining support from many businesses and public interest groups in the United States and seem likely to be more widely utilized in the future.

43. The passage mentions which of the following as a major advantage of domsats?
    (A) They are inexpensive to operate.
    (B) They easily connect distant points.
    (C) They can be directed by remote control.
    (D) They can be built to be very light.

44. According to the passage, the use of domsats is especially valuable for which of the following?
    (A) Mountain areas
    (B) Busy cities
    (C) Small countries
    (D) Private businesses

45. Who objects to the use of domsats?
    (A) Managers of international business groups
    (B) People in small villages
    (C) Operators of conventional communications systems
    (D) Large public interest groups

46. According to the passage, future United States domsats will probably
    (A) be produced competitively
    (B) carry telephone messages only
    (C) become a government monopoly
    (D) increase in use

BEST COPY AVAILABLE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 47-53

Allelomimetic behavior may be defined as behavior in which two or more individual animals do the same thing, with some degree of mutual stimulation and coordination. It can only evolve in species with sense organs that are well enough developed so that continuous sensory contact can be maintained. It is found primarily in vertebrates, in those species that are diurnal, and usually in those that spend much of their lives in the air, in open water, or on open plains.

In birds, allelomimetic behavior is the rule rather than the exception, though it may occasionally be limited to particular seasons of the year as it is in the redwing blackbird. Its principal function is that of providing safety from predators, partly because the flock can rely on many pairs of eyes to watch for enemies, and partly because if one bird reacts to danger, the whole flock is warned.

Among mammals, allelomimetic behavior is very rare in rodents, which almost never move in flocks or herds. Even when they are artificially crowded together, they do not conform in their movements. On the other hand, such behavior is a major system among large hoofed mammals such as sheep.

In the pack-hunting carnivores, allelomimetic behavior has another function, that of cooperative hunting for large prey animals such as moose. Wolves also defend their dens as a group against larger predators such as bears.

Finally, allelomimetic behavior is highly developed among most primate groups, where it has the principal function of providing warning against predators, though combined defensive behavior is also seen in troops of baboons.

47. The main topic of the passage is the
    (A) value of allelomimetic behavior in vertebrate and invertebrate species
    (B) definition and distribution of allelomimetic behavior
    (C) relationship of allelomimetic behavior to the survival of the fittest
    (D) personality factors that determine when an individual animal will show allelomimetic behavior

48. According to the passage, the primary function of allelomimetic behavior in birds is to
    (A) defend nests against predators
    (B) look at each other
    (C) locate prey
    (D) warn others of predators

49. Which of the following places is the most likely setting for allelomimetic behavior?
    (A) A lake
    (B) A cave
    (C) An underground tunnel
    (D) A thick forest

50. According to the passage, what happens to the behavior of rodents when they are artificially crowded together?
    (A) Their allelomimetic behavior increases.
    (B) Continuous cooperation between them is maintained.
    (C) They become aggressive and attack each other.
    (D) They show little allelomimetic behavior.

3 • 3 • 3 • 3 • 3 • 3 • 3

51. The author implies that allelomimetic behavior occurs most often among animals that
    (A) prey on other animals
    (B) are less intelligent than their enemies
    (C) move in groups
    (D) have one sense organ that dominates perception

52. Which of the following is most clearly an example of allelomimetic behavior?
    (A) Bears hunting for carnivores
    (B) Cattle fleeing from a fire
    (C) Horses running at a racetrack
    (D) Dogs working with police officers

53. Which of the following groups of human beings would probably show the greatest amount of allelomimetic behavior?
    (A) A group of students taking a test
    (B) Tennis players competing in a tournament
    (C) A patrol of soldiers scouting for the enemy
    (D) Drivers waiting for a traffic light to change

BEST COPY AVAILABLE

3 • 3 • 3 • 3 • 3 • 3 • 3

Questions 54-60

Criticism of research lays a significant foundation for future investigative work, but when students begin their own projects, they are likely to find that the standards of validity in field work are considerably more rigorous than the standards for most library research. When students are faced with the concrete problem of proof by field demonstration, they usually discover that many of the "important relationships" they may have criticized other researchers for failing to demonstrate are very elusive indeed. They will find, if they submit an outline or questionnaire to their classmates for criticism, that other students make comments similar to some they themselves may have made in discussing previously published research. For example, student researchers are likely to begin with a general question but find themselves forced to narrow its focus. They may learn that questions whose meanings seem perfectly obvious to them are not clearly understood by others, or that questions which seemed entirely objective to them appear to be highly biased to someone else. They usually find that the formulation of good research questions is a much more subtle and frustrating task than is generally believed by those who have not actually attempted it.

54. What does the author think about trying to find weaknesses in other people's research?
    (A) It should only be attempted by experienced researchers.
    (B) It may cause researchers to avoid publishing good work.
    (C) It is currently being done to excess.
    (D) It can be useful in planning future research.

55. According to the passage, what is one major criticism students often make of published research?
    (A) The research has not been written in an interesting way.
    (B) The research has been done in unimportant fields.
    (C) The researchers did not adequately establish the relationships involved.
    (D) The researchers failed to provide an appropriate summary.

56. According to the passage, how do students in class often react to another student's research?
    (A) They react the way they do to any other research.
    (B) They are especially critical of the quality of the research.
    (C) They offer unusually good suggestions for improving the work.
    (D) They show a lot of sympathy for the student researcher.

3 • 3 • 3 • 3 • 3 • 3 • 3

57. According to the passage, what do student researchers often learn when they discuss their work in class?
    (A) Other students rarely have objective comments about it.
    (B) Other students do not believe the researchers did the work themselves.
    (C) Some students feel that the conclusions are too obvious.
    (D) Some students do not understand the meaning of the researchers' questions.

58. According to the passage, student researchers may have to change their research projects because
    (A) their budgets are too high
    (B) their original questions are too broad
    (C) their teachers do not give adequate advice
    (D) their time is very limited

59. What does the author conclude about preparing suitable questions for a research project?
    (A) It is more difficult than the student researcher may realize.
    (B) The researcher should get help from other people.
    (C) The questions should be brief so that they will be understood.
    (D) It is important to follow formulas closely.

60. What does this passage mainly discuss?
    (A) The decreasing emphasis on library research
    (B) How to publish controversial questionnaires
    (C) The role of criticism in new research
    (D) How to submit an outline for criticism

THIS IS THE END OF SECTION 3

IF YOU FINISH BEFORE TIME IS CALLED, CHECK YOUR WORK ON SECTION 3 ONLY.
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.

STOP STOP STOP STOP STOP STOP STOP

BEST COPY AVAILABLE

NOTICE TO TOEFL EXAMINEES

If you took TOEFL on the test date shown below, you may order a copy of the answer key (a list of the correct answers for the questions in the test), a copy of your completed answer sheet, and a cassette recording of Section 1 (listening comprehension). These materials will be available for only a limited time after the test date; this service is subject to change without notice.

HOW TO ORDER

To order these materials, you must complete the order form and mailing slip below. PRINT or TYPE all the required information. The form will be returned to you if you do not give complete information or if you do not enclose the correct fee.

Detach the order form and mail it and the correct payment (or receipt, if you pay by bank draft) to TOEFL, Box 2877, Princeton, NJ 08541, USA. Your materials will be mailed to the address you give on the form.

Materials requested soon after the test date shown on the order form will be mailed about sixty days after that date (they cannot be mailed earlier). Orders received more than sixty days after the test date may take up to thirty additional days to be filled. This order form will not be accepted after:

SEPTEMBER 13, 1983.

FEE

If your mailing address is in the United States (including its territories and possessions), Puerto Rico, or Canada, the fee for the materials is US$10. For addresses in all other countries, the fee is US$12.

DETACH HERE.

Cut Along This Line.

TOEFL Test Materials Order Form

This form may be used only for the following test date:

MAY 14, 1983

Enclosed:
☐ US$10
☐ US$12

Application Number: _____ Test Center Number: _____

Name: _____
Family Name (Surname) First Name Middle Name

Birth Date: / / Signature: _____

Name: _____

Address: _____

_____

_____

Do Not Detach Mailing Slip.

TOEFL
Box 2877
Princeton, NJ 08541, USA

AIR MAIL
875-52

Your Name: _____

Your Address: _____

_____

_____

110

YOUR TOEFL TEST MATERIALS ARE ENCLOSED.

TEST OF ENGLISH AS A FOREIGN LANGUAGE
SCRIPT FOR THE LISTENING COMPREHENSION SECTION

Copyright © 1983 by Educational Testing Service. All rights reserved.
Princeton, N.J. 08541

E. T. S. CONFIDENTIAL

SCRIPT

(MA) In a moment, you are going to hear an introductory statement by the three people who recorded this test. The purpose of this introduction is to give the proctor an opportunity to adjust the recording equipment or make changes in your seating arrangement before the actual test begins. Now listen carefully to the statement by each of the speakers whom you will hear on the test.

(Spoken in turn by MB, W, and MA) Flight number 53 to Paris will depart from gate six at 9:30 p.m. Will all passengers holding tickets kindly proceed to gate six at this time. Thank you. (5 second pause).

Now open your test book. Read the directions in your test book as you listen to the directions on the recording. (8 seconds)

LISTENING COMPREHENSION

In this section of the test, you will have an opportunity to demonstrate your ability to understand spoken English. There are three parts to this section, with special directions for each part.

*Part A

Directions: For each question in Part A, you will hear a short statement. The statements will be spoken just one time. They will not be written out for you, and you must listen carefully to understand what the speaker says.

After you hear a statement, read the four sentences in your test book, marked (A), (B), (C), and (D), and decide which one is closest in meaning to the statement you heard. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen so that the letter inside the oval cannot be seen.

Look at Example I.

You will hear: (W) John is a better student than his brother James.

You will read:
(A) John does better in his studies than James.
(B) James is bigger than his brother John.
(C) John has only one brother.
(D) The teacher likes James better than John.

Sentence (A), "John does better in his studies than James," means most nearly the same as the statement "John is a better student than his brother James." Therefore, you should choose answer (A).

Look at Example II.

You will hear: (MB) The truck traffic on this highway is so heavy I can barely see where I'm going.

You will read:
(A) The traffic isn't bad today.
(B) The trucks weigh a lot.
(C) There are a lot of trucks on the highway.
(D) The highway has been closed to heavy trucks.

Sentence (C), "There are a lot of trucks on the highway," is closest in meaning to the sentence "The truck traffic on this highway is so heavy I can barely see where I'm going." Therefore, you should choose answer (C).

Now let us begin Part A with question number one.

*Unless otherwise noted, all directions will be read by (MA).

1. (MA) After class, go straight to the post office. (12 seconds)
2. (MB) Geometry is hard for me, but algebra is harder. (12 seconds)
3. (MA) Greg thought he could do it himself. (12 seconds)
4. (W) No sooner had he finished washing his car than it started to rain. (12 seconds)
5. (MB) Cooking is not allowed in this dormitory. (12 seconds)
6. (MB) You'd better call Margaret before it gets too late. (12 seconds)
7. (W) He works in the library only at night. (12 seconds)
8. (MA) How about joining us for dinner? (12 seconds)
9. (MB) Jerry hates washing and ironing his own clothes. (12 seconds)
10. (MA) Debbie wrote a check for her son's doctor bill. (12 seconds)
11. (W) Today is Thursday and the swimming pool is supposed to open the day after tomorrow. (12 seconds)
12. (MB) I got my passport photo taken at Nelson Studios. (12 seconds)
13. (W) I was told that you know a lot about this museum. (12 seconds)
14. (MA) Linda's performance wasn't what I'd expected. (12 seconds)

Go on to the next page. (8 seconds)

15. (W) If Dorothy helped you, you'd finish in no time. (12 seconds)
16. (MB) I just ran out of supplies. (12 seconds)
17. (W) This seminar is fascinating, don't you think? (12 seconds)
18. (MB) She didn't sell her old record player after all. (12 seconds)
19. (W) The encyclopedias were out of order. (12 seconds)
20. (W) Our group broke up at two. (12 seconds)

This is the end of Part A. Now look at the directions for Part B as they are read to you.

Part B

Directions: In Part B you will hear short conversations between two speakers. At the end of each conversation, a third voice will ask a question about what was said. The questions will be spoken just one time. After you hear a conversation and the question about it, read the four possible answers in your test book and decide which one is the best answer to the question you heard. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen.

Look at Example I.

You will hear: (MB) Is there any assignment for next Tuesday?

(W) Nothing to read or write. But we're supposed to listen to a radio program and be ready to talk about it in class.

(MA) What have the students been asked to do before Tuesday?

You will read:

(A) Read a book.
(B) Write a composition.
(C) Talk about a problem.
(D) Listen to the radio.

From the conversation you know that the assignment is to listen to a radio program and be ready to talk about it. The best answer, then, is (D), "Listen to the radio." Therefore, you should choose answer (D).

Now let us begin Part B with question twenty-one.

21. (MB) Good morning, may I help you?

(W) Yes. I'd like to cash these traveler's checks first and then open a savings account.

(MA) Where does this conversation probably take place? (12 seconds)

22. (W) We really must go to the new movie in town.

(MB) Let's eat first.

(MA) What does the man want to do? (12 seconds)

23. (MB) I think it's starting to snow.

(W) Starting to snow! The ground's already covered!

(MA) What does the woman mean? (12 seconds)

24. (W) John seems to have lost a lot of weight recently.

(MA) Yes, he's been training hard with the soccer team.

(MB) What has John been doing? (12 seconds)

25. (W) How do you find your new apartment?

(MB) Well, it's quite nice, really, although I'm having a hard time getting used to such a big building.

(MA) What is the man's problem? (12 seconds)

26. (MA) Have you ever put one of these together before?

(MB) No, never, but I think if we carry out these instructions exactly, we won't have any trouble.

(W) What is it important for them to do? (12 seconds)

27. (MB) The front tire is flat and the seat needs to be raised.

(W) Why not take it to Mr. Smith?

(MA) What kind of work does Mr. Smith probably do? (12 seconds)

28. (MA) I haven't decided which color to paint my room—white or yellow.

(W) Isn't easy to choose, is it?

(MB) What does the woman mean? (12 seconds)

29. (MB) If you'd like to take the package with you, Miss, it won't take long to wrap.

(W) There's no rush. Could you please have it delivered this week?

(MA) What does the woman mean? (12 seconds)

30. (W) The map shows that this street goes downtown.

    (MB) Yes, but what we want to know is how to get to the park.

    (MA) What does the man mean? (12 seconds)

31. (MA) My typing isn't dark enough and the paper doesn't look good.

    (W) Why not change the typewriter ribbon and see if that helps?

    (MB) What does the woman advise the man to do? (12 seconds)

32. (W) Didn't you tell Tom about the meeting?

    (MB) Whatever I say to him goes in one ear and out the other.

    (MA) What does the man mean? (12 seconds)

33. (W) You look like you have your hands full. Do you need some help carrying those boxes?

    (MA) I sure do!

    (MB) What will the woman do? (12 seconds)

34. (W) Are you coming with us?

    (MB) No, I have to catch up on my zoology assignments.

    (MA) What does the man mean? (12 seconds)

35. (MA) I heard that the newspaper gave that book a terrible review.

    (W) It depends on which newspaper you read. (pronounce "reed")

    (MB) What does the woman mean? (12 seconds)

This is the end of Part B. Go on to the next page. (8 seconds) Now look at the directions for Part C as they are read to you.

Part C

Directions: In this part of the test, you will hear several short talks and conversations. After each talk or conversation, you will be asked some questions. The talks and questions will be spoken just one time. They will not be written out for you, so you will have to listen carefully to understand what the speaker says.

After you hear a question, read the four possible answers in your test book and decide which one is the best answer to the question you heard. Then, on your answer sheet, find the number of the question and blacken the space that corresponds to the letter of the answer you have chosen.

Listen to this sample talk.

You will hear: (W) Ellis Island is closed now—to all but the tourists, that is. This island, in New York harbor, was once one of the busiest places in America. It was the first stop for all immigrants arriving by ship from Europe, Africa and western Asia. Normally, immigrants came to Ellis Island at the rate of 5,000 a day, but at times twice that many would land in a single day. Most were processed through and ferried to the mainland on the same day. A total of 15 million people came to America by way of Ellis Island. With the advent of air travel, the island fell into disuse. Today it serves only as a reminder to tourists of the heritage of modern America.

Now look at the following example.

You will hear: (MB) How did people generally arrive at Ellis Island?

You will read:
(A) By plane.
(B) By ship.
(C) By train.
(D) By bus.

The best answer to the question "How did people generally arrive at Ellis Island?" is (B), "By ship." Therefore, you should choose answer (B).

Now look at the next example.

You will hear: (MB) Who visits Ellis Island today?

You will read:
(A) New immigrants.
(B) International traders.
(C) Fishermen.
(D) Tourists.

The best answer to the question "Who visits Ellis Island today?" is (D), "Tourists." Therefore, you should choose answer (D).

Now let us begin Part C with question number thirty-six.

(MA) Questions 36-41 refer to the following lecture.

(W) Good morning, students. I hope you have been able to read the two books about speech and hearing problems that I put in the library. Today's lecture deals with the presence of the unusually large deaf population that existed on the Massachusetts island of Martha's Vineyard for about three centuries. From the settlement of the island in the 1640's to the twentieth century, the people there, who were descended from only twenty-five or thirty original families, married mainly other residents of the island. They formed a highly inbred group, producing an excellent example of the genetic patterns for the inheritance of deafness. Indeed, in the late 1800's, one out of every twenty-five people in one village on the island was born deaf, and the island as a whole had a deafness rate at least seventeen times greater than that of the rest of the United States. Even Alexander Graham Bell, the inventor of the telephone and a prominent researcher into hearing loss, visited Martha's Vineyard to study the population, but because the principles of genetics and inheritance were still unknown, he was not able to explain the patterns of deafness, and why a deaf parent did not always have deaf children. In the twentieth century, the local population has mixed with people off the island, and the rate of deafness has fallen.

36. (MB) Where does this talk take place? (12 seconds)

37. (MB) What is unusual about the island of Martha's Vineyard? (12 seconds)

38. (MB) Why were so many people there deaf? (12 seconds)

39. (MB) The island's rate of deafness was how many times greater than that of the rest of the United States? (12 seconds)

40. (MB) What did Alexander Graham Bell hope to do when he went to the island? (12 seconds)

41. (MB) According to the talk, how has the island changed in the twentieth century? (12 seconds)

(MA) Questions 42-46 refer to the following dialogue.

(MB) Have you ever visited a redwood forest? I recently had a chance to go to Muir Woods National Monument north of San Francisco.

(W) I've never seen a redwood tree. I really can't imagine how big they are.

(MB) The coastal redwoods are the tallest living things; some are more than three hundred fifty feet high. But, none of the trees in Muir Woods is that tall. You have to go further north in California to see the tallest trees.

(W) You said that Muir Woods is near San Francisco? I guess it must be quite a tourist attraction.

(MB) Yes. It's less than an hour's drive away, so it's easy to get to.

(W) I've heard that many redwood trees are thousands of years old. Are the ones in Muir Woods that old?

(MB) The oldest documented age for a coastal redwood is more than two thousand years. The trees in Muir Woods are four hundred to eight hundred years old.

(W) Why have they survived so long?

(MB) They have remarkable resistance to forest fires. Their tough, thick bark protects the trees during a fire. The coastal redwoods also like a damp, foggy climate.

(W) Then, since Muir Woods is near foggy San Francisco, it must be ideal for the trees' survival. I can't wait to go there and see them!

42. (MA) What is the main subject of this conversation? (12 seconds)

43. (MA) Where can the tallest trees be found? (12 seconds)

44. (MA) Why do many tourists visit Muir Woods rather than other redwood forests? (12 seconds)

45. (MA) Approximately what is the oldest documented age for a redwood tree? (12 seconds)

46. (MA) What has contributed most to the redwoods' survival? (12 seconds)

(MA) Questions 47-50 are based on the following announcement.

(MB) May I have your attention please? We will be closing in a few minutes. Please return reference books to their shelves. People who wish to check out reserve books for overnight use may do so now.

47. (W) For whom is the announcement primarily intended? (12 seconds)

48. (W) When will the building be closed? (12 seconds)

49. (W) What does the man ask the people to do? (12 seconds)

50. (W) What does the man say about reserve books? (12 seconds)

(MA) Stop work on Section 1.

End of Recording.

APPENDIX B

General Outline of Communication Skills

April 1983

GENERAL OUTLINE OF COMMUNICATION SKILLS

Michael Canale, O.I.S.E., Toronto

(Prepared for the project "Discourse Skills and the TOEFL" directed by Richard Duran, Educational Testing Service)

COMPETENCE AREA	RELEVANT MODE(S)
A. Grammatical competence	
1. Pronunciation:	
1.1. Lexical items in connected speech (at normal rate of speech)	L, S, R (oral)
1.2. Modifications to normal pronunciation of lexical items at word boundaries (e.g., liaison and elision) and in unstressed syllables (e.g., vowel and consonant reduction)	L, S
1.3. Normal word stress in connected speech	L, S, R (oral)
1.4. Emphatic or contrastive word stress (e.g., Mary is happy but Paul is unhappy.)	L, S. R (oral)
1.5. Normal intonation patterns in connected speech (e.g., for imperatives, interrogatives, etc.)	L, S
1.6. Emphatic of contrastive intonation patterns for different clause types (e.g., He has arrived? with rising intonation to signal an interrogative)	L, S. R (oral)
1.7. Normal pauses, loudness, and rate of speech	L, S
1.8. Modifications to normal pauses, loudness, and rate of speech for emphatic or contrastive purposes	L, S
2. Orthography:	
2.1. Graphemes (individually and in sequance)	R, W
2.2. Spelling (including capitalization and diacritics) for individual lexical items	R, W

2.3.	Spelling of compounds (e.g., use of hyphens, as in lion-like, level-headed and vice-president)	R, W
2.4.	Spelling of contractions (e.g., can't)	R, W
2.5.	Spelling of abbreviations (e.g., cont'd)	R, W
2.6.	Spelling of possessive noun forms (e.g., John's)	R, W
2.7.	Common punctuation conventions (e.g., capitalization at beginning of a sentence and use of commas, quotes, etc.)	R, W
2.8.	Conventions for marking emphasis (e.g., underlining, italics, bold-face type, capitalization)	R, W

3. Vocabulary:

3.1.	Literal meaning of common content words, in context, related to academic and social topics	L, S, R, W
3.2.	Literal meaning of common function words in context (e.g., prepositions, articles)	L, S, R, W
3.3.	Meaning of idioms and formulaic expressions in context (e.g., That test was her Little Big Horn; Take care!)	L, S, R, W
3.4.	Extended or figurative meaning of words in context (e.g., metaphorical uses of words as in Marriage is a business partnership)	L, S, R, W
3.5.	Synonyms, antonyms, and homonyms of common content words in context	L, S, R, W

4. Word formation:

4.1.	Inflection, in context, of nouns for number	L, S, R, W
4.2.	Inflection, in context, of demonstrative and possessive adjective for number	L, S, R, W
4.3.	Inflection, in context, of verbs for person, number and tense	L, S, R, W
4.4.	Agreement, in context, of pronouns with nouns	L, S, R, W
4.5.	Agreement, in context, of demonstrative and possessive adjectives with nouns and pronouns	L, S, R, W

4.6. Agreement, in context, of nouns and pronouns with verbs (person and number for verbs, case for pronouns) L, S, R, W

4.7. Derivational relationships (e.g., among attacker and attack as a verb or noun) in context L, S, R, W

4.8. Variation at word boundaries in context (e.g., a and an) L, S, R, W

5. Sentence formation:

5.1. Basic form of common sentence and subsentence structures, in context, relevant to academic and social language-use situation (e.g., subject - verb - complement word order for a simple declarative sentence) L, S, R, W

5.2. Literal meaning of a sentence having a given structure (with vocabulary), in context L, S, R, W

B. Sociolinguistic competence

1. In academic and social situations that vary according to sociolinguistic variables, such as number and status of participants (e.g., peers, strangers, authorities), setting (e.g., formal/informal, public/private, familiar/unfamiliar), channel (e.g., face-to-face, radio, letter, telephone), purpose (e.g., routine/unusual, open-ended/fixed) and amount of shared information:

1.1. Grammatical forms (e.g., pronunciation, etc.) appropriate for different communicative functions, such as supplying or requesting information, persuading, seeking approval, inviting, promising, complaining, socializing L, S, R, W

1.2. Formulaic expressions appropriate for different communicative functions (e.g., Hello/Goodbye on the telephone rather than in written communication) L, S, R, W

1.3. Appropriate grammatical forms for signaling attitudes (e.g., politeness, sincerity, empathy, certainty, anger) L, S, R, W

1.4. Grammatical forms as indicators of social and geographical background (e.g., dialect features) L, S, R, W

C. Discourse competence

1. Cohesion in genres of discourse relevant to academic and social language use:

1.1. Lexical cohesion devices for:

conciseness: e.g., pronouns, synonyms

continuity: e.g., repetition of a vocabulary item

transition: e.g., logical connectors such as however

emphasis: e.g., choice of unexpected vocabulary L, S, R, W

1.2. Grammatical cohesion devices for:

conciseness: e.g., ellipsis

continuity: e.g., parallel structures, lists

transition: e.g., transitional sentences to introduce ideas

emphasis: e.g., focusing structures, such as What is needed is... L, S, R, W

2. Coherence in genres of discourse relevant to academic and social language use:

2.1. Conversational discourse patterns: turn-taking rules (as in a telephone conversation) L, S

2.2. conversational discourse patterns: acceptable oganization of ideas (literal meanings and communicative functions) in conversation in terms of:
development: e.g., sequencing and direction of ideas

continuity: e.g., relevance and consitency of ideas

balance: e.g., treamtent of main vs. supporting ideas

completeness: e.g., thorough discussion of a topic L, S, R, W

2.3. nonconversational discourse patterns: acceptable organization of ideas (literal meanings and communicative functions) in terms of:
development

continuity

balance

completeness L, S, R, W

3. Transposing information in nonverbal/graphic form to and from oral and written discourse (e.g., diagrams, graphs, and tables) L, S, R, W

D. Strategic competences

1. Compensatory strategies for grammatical difficulties:

1.1.	Reference books (e.g., dictionary, grammar book)	R, W
1.2.	Reference centers (e.g., library, resource center), including use of index cards, knowledge of Dewey decimal system	R, W
1.3.	Phonetic spelling as a guide to pronunciation (e.g., International Phonetic Alphabet)	S, R
1.4.	Grammatical and lexical paraphrase (e.g., use of general vocabulary items such as place, person, thing, way followed by a descriptive phrase; use of structures such as ask someone - infinitive rather than demand that - subjunctive)	L, S, R, W
1.5.	Form of requests for repetition, clarification, or slower speech	L, S
1.6.	Use of nonverbal symbols (e.g., gestures, drawings)	L, S, R, W
1.7.	Use of contextual clues for inferences about literal meaning of unfamiliar vocabulary and structures	L, S, R
1.8.	Use of word formation rules to draw inferences about literal meaning of unfamiliar vocabulary and structures (e.g., coinage of fish-house to express aquarium)	L, S, R, W
1.9.	Other (e.g., avoidance of unfamiliar topics, memorization of certain verbal repertoires)	L, S, R, W

2. Compensatory strategies for sociolinguistic difficulties:

2.1.	Single grammatical form for different communicative functions (e.g., a declarative such as Dinner is at 5:00 with varying in tonation to signal a statement, a question, a promise, an order, an invitation--all depending on sociolinguistic context)	L, S, R, W
2.2.	Use of sociolinguistically neutral grammatical forms when uncertain about appropriateness of other forms in a given sociolinguistic context (e.g., in meeting someone, omission of the person's name if unsure about using his or her first name versus title)	S, W
2.3.	Use of first language knowledge about appropriateness of grammatical forms or communicative functions in a given sociolinguistic context	L, S, R, W

2.4.	Use of contextual clues for inferences about social meaning (communicative function, etc.) in unfamiliar sociolinguistic situations or when unfamiliar grammatical forms are used	L, S, R
3.	Compensatory strategies for discourse difficulties:	
3.1.	Use of nonverbal symbols or of emphatic stress and intonation to indicate cohesion and coherence (e.g., use of drawings to indicate sequencing of actions/ideas)	L, S, R, W
3.2.	Use of first language knowledge about oral/written discourse patterns when uncertain about such aspects of discourse in second language	L, S, R, W
3.3.	Use of contextual clues for inferences about patterning of literal and social meanings in unfamiliar discourse	L, S, R, W
4.	Compensatory strategies for performance limitations:	
4.1.	Coping with background noise, interruptions, frequent changes in topic/interlocutors, and other distractions	L, S, R, W
4.2.	Use of pause fillers (e.g., well, you know, my, my) to maintain one's turn in conversation while searching for ideas or grammatical forms or while monitoring them)	L, S
5.	Rhetorical enhancement strategies (noncompensatory):	
5.1.	In oral and written discourse, use of structures and vocabulary for special effect (e.g., use of adverbial phrase preposing, as in Out of the woods came...)	L, S, R, W
5.2.	In oral discourse, use of slow, soft, deliberate speech for special effect	L, S
5.3.	In oral and written discourse, use of literary devices (sentence rhythm, alliteration, literary references)	L, S, R, W

APPENDIX C

Checklist of Competence Areas for
Evaluating TOEFL

Developed by Joyce Penfield
Rutgers University

and Richard Duran
Educational Testing Service

based on Appendix B
and additional source materials

A. GRAMMATICAL COMPETENCE

1. Pronunciation

1.1 Lexical items in connected speech in full phonemic form

1.2 Lexical items modified in connected speech

1.2.1 Vowel reduction/deletion or consonant cluster reduction

1.2.2 Palatalization (e.g., wacha/for "what do you")

1.2.3 Contraction

1.3 Stress and intonation patterns marking neutral, nonemotive reference

1.4 Stress and intonation patterns marking contrastive or emphatic meaning

1.5 Tempo, range/height of pitch, or pauses marking emotive or attitudinal meaning

2. Script (written symbols)

2.1 Letters in sequence

2.1.1 Inherent spelling pattern for words

2.1.2 Modification of inherent spelling pattern of stem/root forms (e.g., "cried" from "cry")

2.2 Graphic symbols used to mark structural groupings

2.2.1 Sentences

2.2.2 Clauses (e.g., to separate subordinate/main clause, a clausal list, or two main changes)

2.2.3 Phrases (e.g., to mark introductory phrases, separate a list of items or dates, places)

2.3 Use of graphic symbols to mark speech-based modifications (e.g., contractions)

2.4 Use of graphic symbols to mark specific concepts

2.4.1 Mathematical concepts or relationships (e.g., 1/300 or $5.20

A. (continued)

2.4.2 Identity (e.g., quotes to identify titles, newspapers, as in: "As Smith notes in 'Passages'....")

2.5 Use of graphic symbols to mark emphatic or contrastive meaning

3. Lexicon (vocabulary)

3.1 Literal use of content words in context

3.2 Idiomatic expressions (i.e., words frozen together into a semantic whole whose meaning cannot be determined by combining the meanings of the parts)

3.2.1 Compound nouns

3.2.2 Compound verbs

3.2.3 Others

3.3 Metaphorical uses of words/expressions (i.e., unique extension of literal reference to nonliteral meaning) (e.g., "the flower of my life")

3.4 Literal use of function words in context

3.4.1 Location; direction (e.g., this/that, near, in)

3.4.2 Specificity

3.4.2.1 Definite

3.4.2.2 Indefinite

3.4.3 Quantity; amount (e.g., many, few)

3.4.4 Time: frequency of occurrence and span (e.g., never, since)

3.4.5 Instrument or means (e.g., he walks with a cane)

3.4.6 Possession (e.g., the wheel of the car)

3.4.7 Comparison/degree (e.g., more intelligent than)

3.4.8 Negation

A. (continued)

4. Morphology (word formation)

4.1 Inflection

4.1.1 Number (e.g., books)

4.1.2 Possession (e.g. John's book)

4.1.3 Person in third person present (e.g., he plays)

4.1.4 Tense

4.1.4.1 Present progressive (e.g. I'm coming)

4.1.4.2 Simple past in regular and irregular form (e.g., he walked/ he sang)

4.1.4.3 Present perfect in regular and irregular forms (e.g. he has walked / he has sung / he has spoken)

4.1.4.4 Past perfect (e.g. he had walked)

4.1.5 Comparative/superlative (e.g., sadder, saddest)

4.2 Derivational relationships in context (e.g., attack/attacker)

5. Sentence formation (ordering of logical constituents)

5.1 Simple sentence word ordering

5.1.1 Declarative, active statements

5.1.2 Questions (yes-no or WH)

5.1.3 Imperatives

5.1.4 Passives (with or without a stated agent)

5.1.5 Existential THERE statements (e.g., There is a rat in the room.)

5.2 Compound sentence word ordering (coordinated main clauses)

5.3 Complex sentence with subordinate clause outside main clause (e.g., When he's ready, he'll call.)

A. (continued)

5.4 Sentences with subordinate clause embedded in main clause

5.4.1 Noun phrase (complementation or nominalization)

5.4.2 Relative clause (possibly signaled by relative pronoun who, which, that, whose)

5.5 Focus shifting, extraposition, or topicalization (e.g., "What I want is that you come." "That I've had plenty of.")

B. SOCIOLINGUISTIC COMPETENCE

1. Factors defining rules of appropriateness for language usage in a given communicative event

1.1 Status/power relationships between participants

1.1.1 Equal role relationship

1.1.2 Subordinate/superordinate role relationship

1.2 Topic

1.2.1 Academic living

1.2.2 Academic content topic

1.2.3 Social naturalness

1.3 Setting/scene

1.3.1 Formal

1.3.2 Informal

1.4 Channel

1.4.1 Spoken

1.4.2 Written

B. (continued)

1.5 Genre (format of communicative event)

1.5.1 Lecture

1.5.2 Conversation

1.5.3 Letter

1.5.4 Note

1.5.5 Announcement

1.5.6 Dated comments

1.5.7 Written passage

1.5.8 Limited functional interaction or segment there from

1.5.9 Statement

2. Language used to accomplish communicative purposes

2.1 Request information

2.1.1 Direct

2.1.2 Indirect

2.2 Persuade

2.2.1 Direct

2.2.2 Indirect

2.3 Seek approval

2.3.1 Direct

2.3.2 Indirect

2.4 Request help or advice

2.4.1 Direct

2.4.2 Indirect

2.5 Promise/assure

2.5.1 Direct

2.5.2 Indirect

B. (continued)

2.6 Invite

2.6.1 Direct

2.6.2 Indirect

2.7 Complain

2.7.1 Direct

2.7.2 Indirect

2.8 Express regret

2.8.1 Direct

2.8.2 Indirect

2.9 Insult

2.9.1 Direct

2.9.2 Indirect

2.10 Make a suggestion

2.10.1 Direct

2.10.2 Indirect

2.11 Warn

2.11.1 Direct

2.11.2 Indirect

2.12 Request permission

2.12.1 Direct

2.12.2 Indirect

2.13 Express impatience, annoyance

2.13.1 Direct

2.13.2 Indirect

B. (continued)

2.14 Express preference

2.14.1 Direct

2.14.2 Indirect

2.15 Advise

2.15.1 Direct

2.15.2 Indirect

2.16 Give an order

2.16.1 Direct

2.16.2 Indirect

2.17 Purchase/transact

2.17.1 Direct

2.17.2 Indirect

2.18 Refuse assistance

2.18.1 Direct

2.18.2 Indirect

2.19 Deny

2.19.1 Direct

2.19.2 Indirect

2.20 Give information

2.20.1 Direct

2.20.2 Indirect

2.21 Other

2.21.1 Direct

2.21.2 Indirect

B. (continued)

3. Formulaic expressions in routine usage for phatic communication (e.g., "Good morning, class. Today's lecture is....")

4. Indirect perspective toward a communicative event (attitudinal or emotive in nature)

4.1 Sarcasm

4.2 Ridicule

4.3 Defeat

4.4 Frustration

4.5 Criticism

4.6 Doubt

5. Linguistic forms or speech modes not typical in broadcast or standard written English

C. DISCOURSE COMPETENCE

1. Cohesion: linguistic devices used to bind text together

1.1 Lexical cohesion devices

1.1.1 Pronoun reference

1.1.2 Synonymy

1.1.3 Word repetition

1.1.4 Endo-centric reference

1.1.5 Sequence markers: first, second

1.1.6 Use of phrasal conjoiners, e.g., and, or, but, for, etc.

C. (continued)

1.2 Sentence level

1.2.1 Conciseness (e.g., ellipsis, clausal reduction)

1.2.2 Continuity: (e.g., parallel structures, lists)

1.2.3 Semantic relationships (e.g., conjunctive adverbs) between clauses

1.2.3.1 Addition

1.2.3.2 Contrast

1.2.3.3 Illustration

1.2.3.4 Similarity

1.2.3.5 Result, conclusion

1.2.3.6 Emphasis

1.2.3.7 Time

1.2.3.8 Place

1.2.3.9 Condition

1.2.4 Emphasis: (e.g., extraposition)

2. Coherence

2.1 Conversational discourse patterns

2.1.1 Initiating the discourse

2.1.2 Maintaining the discourse

2.1.3 Terminating the discourse

2.2 Structuring of ideas in planned texts

2.2.1 Classification (class-inclusion relations used to discuss topic)

C. (continued)

2.2.2 Illustration (concrete examples or ancedote used to explain topic)

2.2.3 Definition (synonymic/metaphonic relations or negative definition used)

2.2.4 Process (event-order)

2.2.4.1 Events narrated according to significance

2.2.4.2 Events narrated according to chronological order

2.2.5 Description (indirect descriptive statements used to establish affect)

2.2.6 Comparison (semantic similarities and differences explored)

2.2.6.1 Alternating order

2.2.6.2 Sequential order

2.2.7 Cause and effect

2.2.8 Factual development (facts chained together by content progression)

APPENDIX D

Summary of Checklist Results

Entries are organized according to section and item type within section. Collapsed entries across item types within a section and across the entire test are also provided. Fractional entries represent the proportion of items manifesting a given characteristic at least once in either their language stimulus, question, or correct response option. Entries in parentheses are the decimal equivalents of fractions.

Competence Areas for Evaluating the TOEFL

A. GRAMMATICAL COMPETENCE

1. Pronunciation

1.1 Lexical items in connected speech in full phonemic form

I: 50/50 (1.00) Sta: 20/20 (1.00) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 50/150 (0.33)

1.2 Lexical items modified in connected speech

1.2.1 Vowel reduction/deletion or consonant cluster reduction

I: 39/50 (0.78) Sta: 14/20 (0.70) D: 12/15 (0.80) M/EC: 11/15 (0.73)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 39/150 (0.26)

1.2.2 Palatalization (e.g. wacha/for "what do you")

I: 1/50 (0.02) Sta: 1/20 (0.05) D: 0/15 (0.00) M/EC: 0/15 (0.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 1/150 (0.01)

1.2.3 Contraction

I: 18/50 (0.36) Sta: 5/20 (0.25) D: 8/15 (0.53) M/EC: 5/15 (0.33)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 18/150 (0.12)

A. (continued)

1.3 Stress and intonation patterns marking neutral, non-emotive reference

I: 47/50 (0.94) Sta: 17/20 (0.85) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 47/150 (0.31)

1.4 Stress and intonation patterns marking contrastive or emphatic meaning

I: 25/50 (0.50) Sta: 10/20 (0.50) D: 8/15 (0.53) M/EC: 7/15 (0.47)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 25/150 (0.17)

1.5 Tempo, range/height of pitch, or pauses marking emotive or attitudinal meaning

I: 48/50 (0.96) Sta: 18/20 (0.90) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 48/150 (0.32)

2. Script (written symbols)

2.1 Letters in sequence

2.1.1 Inherent spelling pattern for words

I: 48/50 (0.96) Sta: 18/20 (0.90) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 40/40 (1.00) Str: 15/15 (1.00) WE: 15/25 (0.60)

III: 60/60 (1.00) V: 30/30 (1.00) RC: 30/30 (1.00)

Total: 148/150 (0.99)

A. (continued)

2.1.2 Modification of inherent spelling pattern of stem/root forms (e.g., "cried" from "cry")

I:	5/50 (0.10)	Sta:	2/20 (0.10)	D:	1/15 (0.07)	M/EC:	2/15 (0.13)
II:	10/40 (0.25)	Str:	3/15 (0.20)	WE:	7/25 (0.28)		
III:	39/60 (0.65)	V:	9/30 (0.30)	RC:	30/30 (1.00)		
Total:	54/150 (0.36)						

2.2 Graphic symbols used to mark structural groupings

2.2.1 Sentences

I:	49/50 (0.98)	Sta:	20/20 (1.00)	D:	15/15 (1.00)	M/EC:	14/15 (0.93)
II:	38/40 (0.95)	Str:	14/15 (0.93)	WE:	24/25 (0.96)		
III:	56/60 (0.97)	V:	28/30 (0.93)	RC:	30/30 (1.00)		
Total:	145/150 (0.97)						

2.2.2 Clauses (e.g., to separate subordinate/main clause, a clausal list, or two main changes)

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	10/40 (0.25)	Str:	8/15 (0.33)	WE:	2/25 (0.08)		
III:	33/60 (0.55)	V:	3/30 (0.10)	RC:	30/30 (1.00)		
Total:	43/150 (0.29)						

2.2.3 Phrases (e.g., to mark introductory phrases, separate a list of items or dates, places)

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	6/40 (0.15)	Str:	1/15 (0.07)	WE:	5/25 (0.20)		
III:	33/60 (0.55)	V:	7/30 (0.23)	RC:	26/30 (0.87)		
Total:	39/150 (0.26)						

A. (continued)

2.3 Use of graphic symbols to mark speech-based modifications (e.g., contractions)

I:	7/50	(0.14)	Sta:	4/20 (0.20)	D:	3/15 (0.20)	M/EC: 0/15 (0.00)
II:	6/40	(0.15)	Str:	1/15 (0.07)	WE:	1/25 (0.04)	
III:	18/60	(0.30)	V:	5/30 (0.17)	RC:	13/30 (0.43)	
Total:	31/150	(0.21)					

2.4 Use of graphic symbols to mark specific concepts

2.4.1 Mathematical concepts or relationships (e.g., 1/300 or $5.20)

I:	1/50	(0.02)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC: 1/15 (0.07)
II:	1/40	(0.02)	Str:	0/15 (0.00)	WE:	1/25 (0.04)	
III:	8/60	(0.13)	V:	0/30 (0.00)	RC:	8/30 (0.27)	
Total:	10/150	(0.07)					

2.4.2 Identity (e.g., quotes to identify titles, newspapers as in: "As Smith notes in 'Passages'....")

I:	3/50	(0.06)	Sta:	3/20 (0.15)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	14/40	(0.35)	Str:	5/15 (0.33)	WE:	9/25 (0.36)	
III:	30/60	(0.50)	V:	11/30 (0.37)	RC:	19/30 (0.63)	
Total:	47/150	(0.31)					

2.5 Use of graphic symbols to mark emphatic or contrastive meaning

I:	0/50	(0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)					

A. (continued)

3. Lexicon (vocabulary)

3.1 Literal use of content words in context

I:	50/50	(1.00)	Sta: 20/20	(1.00)	D: 15/15	(1.00)	M/EC: 15/15 (1.00)
II:	40/40	(1.00)	Str: 15/15	(1.00)	WE: 25/25	(1.00)	
III:	59/60	(0.98)	V: 29/30	(0.97)	RC: 30/30	(1.00)	
Total:	149/150	(0.99)					

3.2 Idiomatic expressions (i.e., words frozen together into a semantic whole whose meaning can not be determined by combining the meanings of the parts)

3.2.1 Compound nouns

I:	26/50	(0.52)	Sta: 5/20	(0.25)	D: 6/15	(0.40)	M/EC: 15/15 (1.00)
II:	17/40	(0.42)	Str: 7/15	(0.47)	WE: 10/25	(0.40)	
III:	44/60	(0.73)	V: 14/30	(0.47)	RC: 30/30	(1.00)	
Total:	87/150	(0.58)					

3.2.2 Compound verbs

I:	20/50	(0.40)	Sta: 4/20	(0.20)	D: 6/15	(0.40)	M/EC: 10/15 (0.67)
II:	5/40	(0.12)	Str: 2/15	(0.13)	WE: 3/25	(0.12)	
III:	21/60	(0.35)	V: 0/30	(0.00)	RC: 2/30	(0.70)	
Total:	46/150	(0.31)					

3.2.3 Others

I:	17/50	(0.34)	Sta: 8/20	(0.40)	D: 3/15	(0.20)	M/EC: 6/15 (0.40)
II:	2/40	(0.05)	Str: 1/15	(0.07)	WE: 1/25	(0.04)	
III:	11/60	(0.18)	V: 4/30	(0.13)	RC: 7/30	(0.23)	
Total:	30/150	(0.20)					

A. (continued)

3.3 Metaphorical uses of words/expressions (i.e., unique extension of literal reference to non-literal meaning) (e.g., "the flower of my life")

I:	1/50	(0.02)	Sta: 0/20	(0.00)	D: 1/15	(0.07)	M/EC: 0/15 (0.00)
II:	1/40	(0.02)	Str: 0/15	(0.00)	WE: 1/25	(0.04)	
III:	0/60	(0.00)	V: 0/30	(0.00)	RC: 0/30	(0.00)	
Total:	2/150	(0.01)					

3.4 Literal use of function words in context

3.4.1 Location; direction (e.g., this/that, near, in)

I:	29/50	(0.58)	Sta: 6/20	(0.30)	D: 8/15	(0.53)	M/EC: 15/15 (1.00)
II:	14/40	(0.35)	Str: 5/15	(0.33)	WE: 9/25	(0.36)	
III:	45/60	(0.75)	V: 15/30	(0.50)	RC: 30/30	(1.00)	
Total:	88/150	(0.59)					

3.4.2 Specificity

3.4.2.1 Definite

I:	43/50	(0.86)	Sta: 13/20	(0.65)	D: 15/15	(1.00)	M/EC: 15/15 (1.00)
II:	38/40	(0.95)	Str: 14/15	(0.93)	WE: 24/25	(0.96)	
III:	52/60	(0.87)	V: 22/30	(0.73)	RC: 30/30	(1.00)	
Total:	130/150	(0.87)					

3.4.2.2 Indefinite

I:	31/50	(0.62)	Sta: 12/20	(0.60)	D: 4/15	(0.27)	M/EC: 15/15 (1.00)
II:	24/40	(0.60)	Str: 10/15	(0.67)	WE: 14/25	(0.56)	
III:	56/60	(0.93)	V: 26/30	(0.87)	RC: 30/30	(1.00)	
Total:	111/150	(0.74)					

A. (continued)

3.4.3 Quantity; amount (e.g., many, few)

I: 21/50 (0.42) Sta: 3/20 (0.15) D: 3/15 (0.20) M/EC: 15/15 (1.00)

II: 15/40 (0.37) Str: 4/15 (0.27) WE: 11/25 (0.44)

III: 40/60 (0.67) V: 10/30 (0.33) RC: 30/30 (1.00)

Total: 76/150 (0.51)

3.4.4 Time: frequency of occurrence and span (e.g., never, since)

I: 20/50 (0.40) Sta: 7/20 (0.35) D: 4/15 (0.27) M/EC: 9/15 (0.60)

II: 4/40 (0.10) Str: 0/15 (0.00) WE: 4/25 (0.12)

III: 34/60 (0.57) V: 4/30 (0.13) RC: 30/30 (1.00)

Total: 58/150 (0.39)

3.4.5 Instrument or means (e.g., he walks <u>with</u> a cane)

I: 0/50 (0.00) Sta: 0/20 (0.00) D: 0/15 (0.00) M/EC: 0/15 (0.00)

II: 3/40 (0.07) Str: 0/15 (0.00) WE: 3/25 (0.12)

III: 5/60 (0.08) V: 1/30 (0.03) RC: 4/30 (0.13)

Total: 8/150 (0.05)

3.4.6 Possession (e.g., the wheel <u>of</u> the car)

I: 23/50 (0.46) Sta: 7/20 (0.35) D: 5/15 (0.33) M/EC: 11/15 (0.73)

II: 15/40 (0.37) Str: 8/15 (0.53) WE: 7/25 (0.28)

III: 32/60 (0.53) V: 2/30 (0.07) RC: 30/30 (1.00)

Total: 70/150 (0.47)

3.4.7 Comparison/degree (e.g., <u>more</u> intelligent <u>than</u>)

I: 12/50 (0.24) Sta: 1/20 (0.05) D: 0/15 (0.00) M/EC: 11/15 (0.73)

II: 5/40 (0.12) Str: 1/15 (0.07) WE: 4/25 (0.11)

III: 13/60 (0.22) V: 3/30 (0.10) RC: 10/30 (0.33)

Total: 30/150 (0.20)

A. (continued)

3.4.8 Negation

I: 26/50 (0.52) Sta: 7/20 (0.35) D: 8/15 (0.53) M/EC: 11/15 (0.73)

II: 1/40 (0.02) Str: 1/15 (0.07) WE: 0/25 (0.00)

III: 28/60 (0.47) V: 2/30 (0.07) RC: 26/30 (0.87)

Total: 55/150 (0.37)

4. Morphology (word formation)

4.1 Inflection

4.1.1 Number (e.g., books)

I: 23/50 (0.46) Sta: 3/20 (0.15) D: 5/15 (0.33) M/EC: 15/15 (1.00)

II: 29/40 (0.72) Str: 11/15 (0.73) WE: 18/25 (0.72)

III: 57/60 (0.95) V: 27/30 (0.90) RC: 30/30 (1.00)

Total: 109/150 (0.73)

4.1.2 Possession (e.g., John's book)

I: 14/50 (0.28) Sta: 2/20 (0.10) D: 1/15 (0.07) M/EC: 11/15 (0.73)

II: 2/40 (0.05) Str: 1/15 (0.07) WE: 1/25 (0.04)

III: 15/60 (0.25) V: 3/30 (0.10) RC: 12/30 (0.40)

Total: 31/150 (0.21)

4.1.3 Person in third person present (e.g., he plays)

I: 33/50 (0.66) Sta: 6/20 (0.30) D: 13/15 (0.87) M/EC: 14/15 (0.93)

II: 16/40 (0.40) Str: 9/15 (0.60) WE: 7/25 (0.28)

III: 30/60 (0.50) V: 7/30 (0.23) RC: 23/30 (0.77)

Total: 79/150 (0.53)

A. (continued)

4.1.4 Tense

4.1.4.1 Present progressive (e.g., I'm coming)

I:	3/50	(0.06)	Sta:	1/20 (0.05)	D:	2/15 (0.13)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)	
III:	4/60	(0.06)	V:	0/30 (0.00)	RC:	4/30 (0.13)	
Total:	7/150	(0.05)					

4.1.4.2 Simple past in regular and irregular form (e.g., he walked/he sang)

I:	24/50	(0.48)	Sta:	11/20 (0.55)	D:	2/15 (0.13)	M/EC: 11/15 (0.73)
II:	16/40	(0.40)	Str:	6/15 (0.40)	WE:	10/25 (0.40)	
III:	29/60	(0.48)	V:	14/30 (0.47)	RC:	15/30 (0.50)	
Total:	69/150	(0.46)					

4.1.4.3 Present perfect in regular and irregular forms (e.g., he has walked/he has sung/ he has spoken)

I:	15/50	(0.30)	Sta:	1/20 (0.05)	D:	3/15 (0.20)	M/EC: 11/15 (0.73)
II:	3/40	(0.07)	Str:	1/15 (0.07)	WE:	2/25 (0.08)	
III:	24/60	(0.40)	V:	1/30 (0.03)	RC:	23/30 (0.92)	
Total:	42/150	(0.28)					

4.1.4.4 Past perfect (e.g., he had walked)

I:	2/50	(0.04)	Sta:	2/20 (0.10)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	1/40	(0.02)	Str:	1/15 (0.07)	WE:	0/25 (0.00)	
III:	8/60	(0.13)	V:	0/30 (0.00)	RC:	8/30 (0.27)	
Total:	11/150	(0.07)					

A. (continued)

4.1.5 Comparative/superlative (e.g., sadder, saddest)

I: 12/50 (0.24) Sta: 1/20 (0.05) D: 0/15 (0.00) M/EC: 11/15 (0.73)

II: 1/40 (0.02) Str: 1/15 (0.07) WE: 0/25 (0.00)

III: 9/60 (0.15) V: 2/30 (0.07) RC: 7/30 (0.23)

Total: 22/150 (0.15)

4.2 Derivational relationships in context (e.g., attack/attacker)

I: 13/50 (0.26) Sta: 1/20 (0.05) D: 0/15 (0.00) M/EC: 11/15 (0.73)

II: 1/40 (0.02) Str: 1/15 (0.07) WE: 1/25 (0.04)

III: 24/60 (0.40) V: 2/30 (0.07) RC: 22/30 (0.88)

Total: 38/150 (0.25)

5. Sentence formation (ordering of logical constituents)

5.1 Simple sentence word ordering

5.1.1 Declarative, active statements

I: 49/50 (0.98) Sta: 19/20 (0.95) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 31/40 (0.77) Str: 11/15 (0.73) WE: 20/25 (0.80)

III: 52/60 (0.87) V: 22/30 (0.73) RC: 30/30 (1.00)

Total: 132/150 (0.88)

5.1.2 Questions (yes-no or WH)

I: 31/50 (0.62) Sta: 1/20 (0.05) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 14/60 (0.23) V: 0/30 (0.00) RC: 14/30 (0.23)

Total: 45/150 (0.30)

A. (continued)

5.1.3 Imperatives

I:	7/50	(0.14)	Sta:	3/20 (0.15)	D:	3/15 (0.20)	M/EC: 1/15 (0.07)
II:	0/40	(0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)	
Total:	7/150	(0.05)					

5.1.4 Passives (with or without a stated agent)

I:	15/50	(0.30)	Sta:	4/20 (0.20)	D:	2/15 (0.13)	M/EC: 9/15 (0.60)
II:	11/40	(0.27)	Str:	7/15 (0.47)	WE:	4/25 (0.11)	
III:	42/60	(0.70)	V:	12/30 (0.40)	RC:	30/30 (1.00)	
Total:	68/150	(0.45)					

5.1.5 Existential THERE statements (e.g., There is a rat in the room.)

I:	1/50	(0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC: 6/15 (0.40)
II:	2/40	(0.05)	Str:	1/15 (0.07)	WE:	1/25 (0.04)	
III:	8/60	(0.13)	V:	0/30 (0.00)	RC:	8/30 (0.63)	
Total:	11/150	(0.07)					

5.2 Compound sentence word ordering (coordinated main clauses)

I:	7/50	(0.14)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC: 6/15 (0.40)
II:	3/40	(0.07)	Str:	2/15 (0.13)	WE:	1/25 (0.04)	
III:	19/60	(0.32)	V:	0/30 (0.00)	RC:	19/30 (0.63)	
Total:	29/150	(0.19)					

5.3 Complex sentence with subordinate clause outside main clause (e.g., <u>When he's ready</u>, he'll call.)

I:	16/50	(0.32)	Sta:	2/20 (0.10)	D:	2/15 (0.13)	M/EC: 12/15 (0.80)
II:	4/40	(0.10)	Str:	2/15 (0.13)	WE:	2/25 (0.08)	
III:	33/60	(0.55)	V:	3/30 (0.10)	RC:	30/30 (1.00)	
Total:	53/150	(0.35)					

A. (continued)

5.4 Sentences with subordinate clause embedded in main clause

5.4.1 Noun phrase (complementation or nominalization)

I:	3/50	(0.06)	Sta:	1/20 (0.05)	D:	2/15 (0.13)	M/EC: 0/15 (0.00)
II:	2/40	(0.05)	Str:	1/15 (0.07)	WE:	1/25 (0.04)	
III:	3/60	(0.05)	V:	1/30 (0.03)	RC:	2/30 (0.07)	
Total:	8/150	(0.05)					

5.4.2 Relative clause (possibly signalled by relative pronouns who, which, that, whose)

I:	14/50	(0.28)	Sta:	1/20 (0.05)	D:	3/15 (0.20)	M/EC: 10/15 (0.67)
II:	5/40	(0.12)	Str:	2/15 (0.13)	WE:	3/25 (0.12)	
III:	32/60	(0.53)	V:	4/30 (0.13)	RC:	28/30 (0.93)	
Total:	51/150	(0.34)					

5.5 Focus shifting, extraposition or topicalization (e.g., "What I want is that you come." "That I've had plenty of.")

I:	2/50	(0.04)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC: 1/15 (0.07)
II:	0/40	(0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)	
III:	11/60	(0.18)	V:	1/30 (0.03)	RC:	10/30 (0.33)	
Total:	13/150	(0.09)					

B. SOCIOLINGUISTIC COMPETENCE

1. Factors defining rules of appropriateness for language usage in a given communicative event

1.1 Status/power relationships between participants

1.1.1 Equal role relationship

I:	11/50	(0.22)	Sta:	0/20 (0.00)	D:	6/15 (0.40)	M/EC: 5/15 (0.33)
II:	0/40	(0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)	
Total:	11/150	(0.07)					

B. (continued)

1.1.2 Subordinate/superordinate role relationship

I: 13/50 (0.26) Sta: 1/20 (0.05) D: 2/15 (0.13) M/EC: 10/15 (0.67)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 13/150 (0.09)

1.2 Topic

1.2.1 Academic living

I: $\overline{X}$ = 1.70 Sta: $\overline{x}$ = 1.50 D: $\overline{x}$ = 1.27 M/EC: $\overline{x}$ = 2.33

II: $\overline{X}$ = 1.52 Str: $\overline{x}$ = 1.53 WE: $\overline{x}$ = 1.52

III: $\overline{X}$ = 1.98 V: $\overline{x}$ = 1.97 RC: $\overline{x}$ = 2.00

1.2.2 Academic content topic

I: $\overline{X}$ = 1.28 Sta: $\overline{x}$ = 1.05 D: $\overline{x}$ = 1.00 M/EC: $\overline{x}$ = 1.80

II: $\overline{X}$ = 1.52 Str: $\overline{x}$ = 1.53 WE: $\overline{x}$ = 1.52

III: $\overline{X}$ = 1.95 V: $\overline{x}$ = 1.90 RC: $\overline{x}$ = 2.00

1.2.3 Social naturalness

I: $\overline{X}$ = 2.45 Sta: $\overline{x}$ = 2.55 D: $\overline{x}$ = 2.40 M/EC: $\overline{x}$ = 2.40

1.3 Setting/scene

1.3.1 Formal

I: 14/50 (0.28) Sta: 0/20 (0.00) D: 4/15 (0.27) M/EC: 10/15 (0.67)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 14/150 (0.09)

B. (continued)

1.3.2 Informal

I: 17/50 (0.34) Sta: 5/20 (0.25) D: 7/15 (0.47) M/EC: 5/15 (0.33)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 17/150 (0.11)

1.4 Channel

1.4.1 Spoken

I: 50/50 (1.00) Sta: 20/20 (1.00) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 50/150 (0.33)

1.4.2 Written

I: 46/50 (0.92) Sta: 16/20 (0.80) D: 15/15 (1.00) M/EC: 15/15 (1.00)

II: 40/40 (1.00) Str: 15/15 (1.00) WE: 25/25 (1.00)

III: 60/60 (1.00) V: 30/30 (1.00) RC: 30/30 (1.00)

Total: 146/150 (0.97)

1.5 Genre (format of communicative event)

1.5.1 Lecture

I: 6/50 (0.12) Sta: 0/20 (0.00) D: 0/15 (0.00) M/EC: 6/15 (0.40)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 6/150 (0.04)

B. (continued)

1.5.2 Conversation

I:	19/50 (0.38)	Sta:	0/20 (0.00)	D:	14/15 (0.93)	M/EC:	5/15 (0.33)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	19/150 (0.13)						

1.5.3 Letter

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

1.5.4 Note

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

1.5.5 Announcement

I:	5/50 (0.10)	Sta:	1/20 (0.05)	D:	0/15 (0.00)	M/EC:	4/15 (0.27)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	5/150 (0.03)						

1.5.6 Dated comments

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

B. (continued)

1.5.7 Written passage

I: 0/50 (0.00) Sta: 0/20 (0.00) D: 0/15 (0.00) M/EC: 0/15 (0.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 30/60 (0.50) V: 0/30 (0.00) RC: 30/30 (1.00)

Total: 30/150 (0.20)

1.5.8 Limited functional interaction or segment there from

I: 34/50 (0.68) Sta: 5/20 (0.25) D: 15/15 (1.00) M/EC: 14/15 (0.93)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 30/60 (0.50) V: 0/30 (0.00) RC: 30/30 (1.00)

Total: 64/150 (0.43)

1.5.9 Statement

I: 34/50 (0.68) Sta: 16/20 (0.80) D: 4/15 (0.27) M/EC: 14/15 (0.93)

II: 36/40 (0.90) Str: 15/15 (1.00) WE: 21/25 (0.84)

III: 37/60 (0.62) V: 30/30 (1.00) RC: 7/30 (0.23)

Total: 107/150 (0.71)

2. Language used to accomplish communicative purposes

2.1 Request information

2.1.1 Direct

I: 29/50 (0.58) Sta: 0/20 (0.00) D: 15/15 (1.00) M/EC: 14/15 (0.93)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 29/60 (0.48) V: 0/30 (0.00) RC: 29/30 (0.97)

Total: 58/150 (0.39)

B. (continued)

2.1.2 Indirect

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

2.2 Persuade

2.2.1 Direct

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	2/15 (0.13)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

2.2.2 Indirect

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

2.3 Seek approval

2.3.1 Direct

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

B. (continued)

2.3.2 Indirect

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	1/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2.4 Request help or advice

2.4.1 Direct

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2.4.2 Indirect

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

2.5 Promise/assure

2.5.1 Direct

I:	2/50 (0.04)	Sta:	1/20 (0.05)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

B. (continued)

2.5.2 Indirect

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2.6 Invite

2.6.1 Direct

I:	1/50 (0.02)	Sta:	1/20 (0.05)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2.6.2 Indirect

I:	2/50 (0.02)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

2.7 Complain

2.7.1 Direct

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

B. (continued)

2.7.2 Indirect

I:	3/50 (0.06)	Sta:	0/20 (0.00)	D:	3/15 (0.20)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	3/150 (0.02)						

2.8 Express regret

2.8.1 Direct

I:	1/50 (0.02)	Sta:	1/20 (0.05)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2.8.2 Indirect

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

2.9 Insult

2.9.1 Direct

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

B. (continued)

2.9.2 Indirect

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2.10 Make a suggestion

2.10.1 Direct

I:	3/50 (0.06)	Sta:	0/20 (0.00)	D:	3/15 (0.20)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	3/150 (0.02)						

2.10.2 Indirect

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

2.11 Warn

2.11.1 Direct

I:	2/50 (0.04)	Sta:	2/20 (0.10)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

B. (continued)

2.11.2 Indirect

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.12 Request permission

2.12.1 Direct

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.12.2 Indirect

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.13 Express impatience, annoyance

2.13.1 Direct

I:	2/50	(0.04)	Sta:	0/20	(0.00)	D:	2/15 (0.13)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	2/150	(0.01)						

B. (continued)

2.13.2 Indirect

I:	2/50	(0.04)	Sta:	0/20	(0.00)	D:	2/15 (0.13)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	2/150	(0.01)						

2.14 Express preference

2.14.1 Direct

I:	1/50	(0.02)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 1/15 (0.07)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	1/150	(0.01)						

2.14.2 Indirect

I:	1/50	(0.02)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 1/15 (0.07)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	1/150	(0.01)						

2.15 Advise

2.15.1 Direct

I:	3/50	(0.06)	Sta:	2/20	(0.10)	D:	1/15 (0.07)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	3/150	(0.02)						

B. (continued)

2.15.2 Indirect

I:	2/50	(0.04)	Sta:	1/20	(0.05)	D:	0/15 (0.00)	M/EC: 1/15 (0.07)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	2/150	(0.01)						

2.16 Give an order

2.16.1 Direct

I:	1/50	(0.02)	Sta:	1/20	(0.05)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	1/150	(0.01)						

2.16.2 Indirect

I:	4/50	(0.08)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 4/15 (0.27)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	4/150	(0.03)						

2.17 Purchase/transact

2.17.1 Direct

I:	2/50	(0.04)	Sta:	0/20	(0.00)	D:	2/15 (0.13)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	2/150	(0.01)						

B. (continued)

2.17.2 Indirect

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

2.18 Refuse assistance

2.18.1 Direct

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

2.18.2 Indirect

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

2.19 Deny

2.19.1 Direct

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

B. (continued)

2.19.2 Indirect

I: 0/50 (0.00) Sta: 0/20 (0.00) D: 0/15 (0.00) M/EC: 0/15 (0.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 0/150 (0.00)

2.20 Give information

2.20.1 Direct

I: 39/50 (0.90) Sta: 11/20 (0.55) D: 13/15 (0.87) M/EC: 15/15 (1.00)

II: 40/40 (1.00) Str: 15/15 (1.00) WE: 25/25 (1.00)

III: 60/60 (1.00) V: 30/30 (1.00) RC: 30/30 (1.00)

Total: 145/150 (0.93)

2.20.2 Indirect

I: 2/50 (0.04) Sta: 0/20 (0.00) D: 2/15 (0.13) M/EC: 0/15 (0.00)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 2/150 (0.01)

2.21 Other

2.21.1 Direct

I: 7/50 (0.14) Sta: 0/20 (0.00) D: 2/15 (0.13) M/EC: 5/15 (0.33)

II: 0/40 (0.00) Str: 0/15 (0.00) WE: 0/25 (0.00)

III: 0/60 (0.00) V: 0/30 (0.00) RC: 0/30 (0.00)

Total: 7/150 (0.05)

B. (continued)

2.21.2 Indirect

I:	5/50 (0.10)	Sta:	0/20 (0.00)	D:	5/15 (0.33)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	5/150 (0.03)						

3. Formulaic expressions in routine usage for phatic communication (e.g.,"Good morning, class. Today's lecture is....)

I:	14/50 (0.93)	Sta:	0/20 (0.00)	D:	4/15 (0.27)	M/EC:	10/15 (0.67)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	14/150 (0.09)						

4. Indirect perspective toward a communicative event (attitudinal or emotive in nature)

4.1 Sarcasm

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

4.2 Ridicule

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

B. (continued)

4.3 Defeat

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

4.4 Frustration

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

4.5 Criticism

I:	2/50 (0.04)	Sta:	0/20 (0.00)	D:	2/15 (0.13)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	2/150 (0.01)						

4.6 Doubt

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

5. Linguistic forms or speech modes <u>not</u> typical in broadcast or standard written English

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	1/15 (0.07)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

C. DISCOURSE COMPETENCE

1. Cohesion: linguistic devices used to bind text together

1.1 Lexical cohesion devices

1.1.1 Pronoun reference

I: 41/50 (0.82) Sta: 15/20 (0.75) D: 13/15 (0.87) M/EC: 13/15 (0.87)

II: 11/40 (0.27) Str: 3/15 (0.20) WE: 8/25 (0.32)

III: 37/60 (0.62) V: 7/30 (0.28) RC: 30/30 (1.00)

Total: 89/150 (0.59)

1.1.2 Synonymy

I: 12/50 (0.24) Sta: 0/20 (0.00) D: 1/15 (0.07) M/EC: 11/15 (0.73)

II: 5/40 (0.12) Str: 2/15 (0.13) WE: 3/25 (0.12)

III: 33/60 (0.55) V: 3/30 (0.10) RC: 30/30 (1.00)

Total: 50/150 (0.33)

1.1.3 Word repetition

I: 13/50 (0.26) Sta: 0/20 (0.00) D: 2/15 (0.13) M/EC: 11/15 (0.73)

II: 5/40 (0.12) Str: 3/15 (0.20) WE: 2/25 (0.08)

III: 32/60 (0.53) V: 2/30 (0.07) RC: 30/30 (1.00)

Total: 60/150 (0.40)

1.1.4 Endo-centric reference

I: 4/50 (0.08) Sta: 0/20 (0.00) D: 3/15 (0.20) M/EC: 1/15 (0.07)

II: 1/40 (0.02) Str: 1/15 (0.07) WE: 0/25 (0.00)

III: 16/60 (0.27) V: 1/30 (0.03) RC: 15/30 (0.50)

Total: 21/150 (0.14)

C. (continued)

1.1.5 Sequence markers: first, second

I: 7/50 (0.14) Sta: 0/20 (0.00) D: 7/15 (0.47) M/EC: 0/15 (0.00)

II: 2/40 (0.05) Str: 2/15 (0.13) WE: 0/25 (0.00)

III: 16/60 (0.27) V: 1/30 (0.03) RC: 15/30 (0.50)

Total: 25/150 (0.17)

1.1.6 Use of phrasal conjoiners (e.g., and, or, but, for, etc.)

I: 12/50 (0.24) Sta: 1/20 (0.05) D: 0/15 (0.00) M/EC: 11/15 (0.73)

II: 11/40 (0.27) Str: 1/15 (0.07) WE: 10/25 (0.40)

III: 35/60 (0.58) V: 5/30 (0.17) RC: 30/30 (1.00)

Total: 58/150 (0.39)

1.2 Sentence level

1.2.1 Conciseness (e.g., ellipsis; clausal reduction)

I: 24/50 (0.48) Sta: 2/20 (0.10) D: 8/15 (0.53) M/EC: 14/15 (0.93)

II: 12/40 (0.30) Str: 6/15 (0.40) WE: 6/25 (0.24)

III: 38/60 (0.63) V: 8/30 (0.27) RC: 30/30 (1.00)

Total: 49/150 (0.74)

1.2.2 Continuity (e.g., parallel structures; lists)

I: 10/50 (0.20) Sta: 0/20 (0.00) D: 0/15 (0.00) M/EC: 10/15 (0.67)

II: 6/40 (0.15) Str: 2/15 (0.13) WE: 4/25 (0.16)

III: 30/60 (0.50) V: 6/30 (0.20) RC: 30/30 (1.00)

Total: 46/150 (0.31)

C. (continued)

1.2.3 Semantic relationships (e.g., conjunctive adverbs) between clauses

1.2.3.1 Addition

I:	10/50 (0.20)	Sta:	1/20 (0.05)	D:	3/15 (0.20)	M/EC:	6/15 (0.40)
II:	3/40 (0.07)	Str:	2/15 (0.13)	WE:	1/25 (0.04)		
III:	14/60 (0.23)	V:	2/30 (0.07)	RC:	12/30 (0.40)		
Total:	27/150 (0.18)						

1.2.3.2 Contrast

I:	14/50 (0.28)	Sta:	1/20 (0.05)	D:	2/15 (0.13)	M/EC:	11/15 (0.73)
II:	2/40 (0.05)	Str:	0/15 (0.00)	WE:	2/25 (0.08)		
III:	26/60 (0.43)	V:	0/30 (0.00)	RC:	26/30 (0.87)		
Total:	42/150 (0.28)						

1.2.3.3 Illustration

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

1.2.3.4 Similarity

I:	0/50 (0.00)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	0/15 (0.00)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	0/150 (0.00)						

C. (continued)

1.2.3.5 Result; conclusion

I:	11/50	(0.22)	Sta:	0/20	(0.00)	D: 0/15 (0.00)	M/EC: 11/15 (0.73)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE: 0/25 (0.00)	
III:	21/60	(0.35)	V:	2/30	(0.07)	RC: 19/30 (0.63)	
Total:	32/150	(0.21)					

1.2.3.6 Emphasis

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D: 0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE: 0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC: 0/30 (0.00)	
Total:	0/150	(0.00)					

1.2.3.7 Time

I:	4/50	(0.08)	Sta:	3/20	(0.15)	D: 0/15 (0.00)	M/EC: 1/15 (0.07)
II:	4/40	(0.10)	Str:	2/15	(0.13)	WE: 2/25 (0.08)	
III:	22/60	(0.37)	V:	2/30	(0.07)	RC: 20/30 (0.67)	
Total:	30/150	(0.20)					

1.2.3.8 Place

I:	1/50	(0.02)	Sta:	0/20	(0.00)	D: 0/15 (0.00)	M/EC: 1/15 (0.07)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE: 0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC: 0/30 (0.00)	
Total:	1/150	(0.01)					

1.2.3.9 Condition

I:	4/50	(0.08)	Sta:	1/2	(0.05)	D: 3/15 (0.20)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE: 0/25 (0.00)	
III:	17/60	(0.28)	V:	1/30	(0.03)	RC: 16/30 (0.53)	
Total:	0/150	(0.00)					

C. (continued)

1.2.4 Emphasis (e.g., extraposition)

I:	1/50 (0.02)	Sta:	0/20 (0.00)	D:	0/15 (0.00)	M/EC:	1/15 (0.07)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	1/150 (0.01)						

2. Coherence

2.1 Conversational discourse patterns

2.1.1 Initiating the discourse

I:	17/50 (0.34)	Sta:	1/20 (0.05)	D:	11/15 (0.73)	M/EC:	5/15 (0.33)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	17/150 (0.11)						

2.1.2 Maintaining the discourse

I:	20/50 (0.40)	Sta:	0/20 (0.00)	D:	15/15 (1.00)	M/EC:	5/15 (0.33)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	20/150 (0.13)						

2.1.3 Terminating the discourse

I:	5/50 (0.10)	Sta:	0/20 (0.00)	D:	0/00 (0.00)	M/EC:	5/15 (0.33)
II:	0/40 (0.00)	Str:	0/15 (0.00)	WE:	0/25 (0.00)		
III:	0/60 (0.00)	V:	0/30 (0.00)	RC:	0/30 (0.00)		
Total:	5/150 (0.03)						

C. (continued)

2.2 Structuring of ideas in planned texts

2.2.1 Classification (class-inclusion relations used to discuss topic)

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.2.2 Illustration (concrete examples or ancedote used to explain topic)

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.2.3 Definition (synonymic/metaphonic relations or negative definition used)

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.2.4 Process (event-order)

2.2.4.1 Events narrated according to significance

I:	6/50	(0.12)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 6/15 (0.40)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	12/60	(0.20)	V:	0/30	(0.00)	RC:	12/30 (0.40)	
Total:	18/150	(0.12)						

C. (continued)

2.2.4.2 Events narrated according to chronological order

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.2.5 Description (indirect descriptive statements used to establish affect)

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.2.6 Comparison (semantic similarities and differences explored)

2.2.6.1 Alternating order

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15 (0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	0/150	(0.00)						

2.2.6.2 Sequential order

I:	1/50	(0.02)	Sta:	0/20	(0.00)	D:	1/15 (0.07)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25 (0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30 (0.00)	
Total:	1/150	(0.01)						

C. (continued)

2.2.7 Cause and effect

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15	(0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25	(0.00)	
III:	0/60	(0.00)	V:	0/30	(0.00)	RC:	0/30	(0.00)	
Total:	0/150	(0.00)							

2.2.8 Factual development (facts chained together by content progression)

I:	0/50	(0.00)	Sta:	0/20	(0.00)	D:	0/15	(0.00)	M/EC: 0/15 (0.00)
II:	0/40	(0.00)	Str:	0/15	(0.00)	WE:	0/25	(0.00)	
III:	18/60	(0.30)	V:	0/30	(0.00)	RC:	18/30	(0.60)	
Total:	18/150	(0.12)							

APPENDIX E

Interagency Language Roundtable
Skill Level Descriptions

INTERAGENCY LANGUAGE ROUNDTABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
SPEAKING

Preface

The following descriptions of proficiency levels 0, 1, 2, 3, 4, and 5 characterize spoken-language use. Each higher level implies control of the previous levels' functions and accuracy. The designation 0+, 1+, 2+, etc. will be assigned when proficiency substantially exceeds one skill level and does not fully meet the criteria for the next level. The "plus-level" descriptions, therefore, are subsidiary to the "base-level" descriptions.

A skill level is assigned to a person through an authorized language examination. Examiners assign a level on a variety of performance criteria exemplified in the descriptive statements. Therefore, the examples given here illustrate, but do not exhaustively describe, either the skills a person may possess or situations in which he/she may function effectively.

Statements describing accuracy refer to typical stages in the development of competence in the most commonly taught languages in formal training programs. In other languages, emerging competence parallels these characterizations, but often with different details.

Unless otherwise specified, the term "native speaker" refers to native speakers of a standard dialect.

"Well-educated," in the context of these proficiency descriptions, does not necessarily imply formal higher education. However, in cultures where formal higher education is common, the language-use abilities of persons who have had such education is considered the standard. That is, such a person meets contemporary expectations for the formal, careful style of the language, as well as a range of less formal varieties of the language.

These descriptions may be further specified by individual agencies to characterize those aspects of language-use performance which are of insufficient generality to be included here.

S-0 NO PROFICIENCY

Unable to function in the spoken language. Oral production is limited to occasional isolated words. Has essentially no communicative ability.

S-0+ MEMORIZED PROFICIENCY

Able to satisfy immediate needs using rehearsed utterances. Shows little real autonomy of expression, flexibility, or spontaneity. Can ask questions or make statements with reasonable accuracy only with memorized utterances or formulae. Attempts at creating speech are usually unsuccessful.

Examples: The S-0+'s vocabulary is usually limited to areas of immediate survival needs. Most utterances are telegraphic; that is, functors (linking words, markers, and the like) are omitted, confused, or distorted. An S-0+ can usually differentiate most significant sounds when produced in isolation, but, when combined in words or groups of words, errors may be frequent. Even with repetition, communication is severely limited even with persons used to dealing with foreigners. Stress, intonation, tone, etc. are usually quite faulty.

S-1 ELEMENTARY PROFICIENCY
(Base Level)

Able to satisfy minimum courtesy requirements and maintain very simple face-to-face conversations on familiar topics. A native speaker must often use slowed speech, repetition, paraphrase, or a combination of these to be understood by an S-1. Similarly, the native speaker must strain and employ real-world knowledge to understand even simple statements/questions from the S-1. An S-1 speaker has a functional, but limited proficiency. Misunderstandings are frequent, but the S-1 is able to ask for help and to verify comprehension of native speech in face-to-face interaction. The S-1 is unable to produce continuous discourse except with rehearsed material.

Examples. Structural accuracy is likely to be random or severely limited. Time concepts are vague. Vocabulary is inaccurate, and its range is very narrow. The S-1 often speaks with great difficulty. By repeating, such speakers can make themselves understood to native speakers who are in regular contact with foreigners but there is little precision in the information conveyed. Needs, experience, or training may vary greatly from individual to individual; for example, S-1s may have encountered quite different vocabulary areas. However, the S-1 can typically satisfy predictable, simple, personal and accommodation needs; can generally meet courtesy, introduction, and identification requirements; exchange greetings; elicit and provide, for example, predictable and skeletal biographical information. An S-1 might give information about business hours, explain routine procedures in a limited way, and state in a simple manner what actions will be taken. The S-1 is able to formulate some questions even in languages with complicated question constructions. Almost every utterance may be characterized by structural errors and errors in basic grammatical relations. Vocabulary is extremely limited and characteristically does not include modifiers. Pronunciation, stress, and intonation are generally poor, often heavily influenced by another language. Use of structure and vocabulary is highly imprecise.

S-1+ ELEMENTARY PROFICIENCY (Higher Level)

Can initiate and maintain predictable face-to-face conversations and satisfy limited social demands. The S-1+ may, however, have little understanding of the social conventions of conversation. The interlocutor is generally required to strain and employ real-world knowledge to understand even some simple speech. An S-1+ may hesitate and may have to change subjects due to lack of language resources. Range and control of the language are limited. Speech largely consists of a series of short, discrete utterances.

Examples: An S-1+ is able to satisfy most travel and accommodation needs and a limited range of social demands beyond exchanges of skeletal biographic information. Speaking ability may extend beyond immediate survival needs. Accuracy in basic grammatical relations is evident, although not consistent. May exhibit the commoner forms of verb tenses, for example, but may make frequent errors in formation and selection. While some structures are established, errors occur in more complex patterns. The S-1+ typically cannot sustain coherent structures in longer utterances or unfamiliar situations. Ability to describe and give precise information is limited. Person, space, and time references are often used incorrectly. Pronunciation is understandable to natives used to dealing with foreigners. Can combine most significant sounds with reasonable comprehensibility, but has difficulty in producing certain sounds in certain positions or in certain combinations. Speech will usually be labored. Frequently has to repeat utterances to be understood by the general public.

S-2 LIMITED WORKING PROFICIENCY (Base Level)

Able to satisfy routine social demands and limited work requirements. Can handle routine work-related interactions that are limited in scope. In more complex and sophisticated work-related tasks, language usage generally disturbs the native speaker. Can handle with confidence, but not with facility, most normal, high-frequency social conversational situations including extensive, but casual conversations about current events, as well as work, family, and autobiographical information. The S-2 can get the gist of most everyday conversations but has some difficulty understanding native speakers in situations that require specialized or sophisticated knowledge. The S-2's utterances are minimally cohesive. Linguistic structure is usually not very elaborate and not thoroughly controlled; errors are frequent. Vocabulary use is appropriate for high-frequency utterances, but unusual or imprecise elsewhere.

Examples: While these interactions will vary widely from individual to individual, an S-2 can typically ask and answer predictable questions in the workplace and give straightforward instructions to subordinates. Additionally, the S-2 can participate in personal and accommodation-type interactions with elaboration and facility; that is, can give and understand complicated, detailed, and extensive directions and make non-routine changes in travel and accommodation arrangements. Simple structures and basic grammatical relations are typically controlled; however, there are areas of weakness. In the commonly taught languages, these may be simple markings such as plurals, articles, linking words, and negatives or more complex structures such as tense/aspect usage, case morphology, passive constructions, word order, and embedding.

S-2+ LIMITED WORKING PROFICIENCY (Higher Level)

Able to satisfy most work requirements with language usage that is often, but not always, acceptable and effective. An S-2+ shows considerable ability to communicate effectively on topics relating to particular interests and special fields of competence. Often shows a high degree of fluency and ease of speech, yet when under tension or pressure, the ability to use the language effectively may deteriorate. Comprehension of normal native speech is typically nearly complete. An S-2+ may miss cultural and local references and may require a native speaker to adjust to his/her limitations in some ways. Native speakers often perceive the S-2+'s speech to contain awkward or inaccurate phrasing of ideas, mistaken time, space, and person references, or to be in some way inappropriate, if not strictly incorrect.

Examples: Typically an S-2+ can participate in most social, formal, and informal interactions; but limitations either in range of contexts, types of tasks, or level of accuracy hinder effectiveness. The S-2+ may be ill at ease with the use of the language either in social interaction or in speaking at length in professional contexts. An S-2+ is generally strong in either structural precision or vocabulary, but not in both. Weakness or unevenness in one of the foregoing, or in pronunciation, occasionally results in miscommunication. Normally controls, but cannot always easily produce general vocabulary. Discourse is often incohesive.

S-3 GENERAL PROFESSIONAL PROFICIENCY (Base Level)

Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Nevertheless, an S-3's limitations generally restrict the professional contexts of language use to matters of shared knowledge and/or international convention. Discourse is cohesive. An S-3 uses the language acceptably, but with some noticeable imperfections; yet, errors virtually never interfere with understanding and rarely disturb the native speaker. An S-3 can effectively combine structure and vocabulary to convey his/her meaning accurately. An S-3 speaks readily and fills pauses suitably. In face-to-face conversation with natives speaking the standard dialect at a normal rate of speech, comprehension is quite complete. Although cultural references, proverbs, and the implications of

nuances and idiom may not be fully understood, the S-3 can easily repair the conversation. Pronunciation may be obviously foreign. Individual sounds are accurate; but stress, intonation, and pitch control may be faulty.

Examples: Can typically discuss particular interests and special fields of competence with reasonable ease. Can use the language as part of normal professional duties such as answering objections, clarifying points, justifying decisions, understanding the essence of challenges, stating and defending policy, conducting meetings, delivering briefings, or other extended and elaborate informative monologues. Can reliably elicit information and informed opinion from native speakers. Structural inaccuracy is rarely the major cause of misunderstanding. Use of structural devices is flexible and elaborate. Without searching for words or phrases, an S-3 uses the language clearly and relatively naturally to elaborate concepts freely and make ideas easily understandable to native speakers. Errors occur in low-frequency and highly complex structures.

S-3+ GENERAL PROFESSIONAL PROFICIENCY (Higher Level)

Is often able to use the language to satisfy professional needs in a wide range of sophisticated and demanding tasks.

Examples: Despite obvious strengths, may exhibit some hesitancy, uncertainty, effort, or errors which limit the range of language-use tasks that can be reliably performed. Typically there is particular strength in fluency and one or more, but not all, of the following: has breadth of lexicon, including low- and medium-frequency items, especially socio-linguistic/cultural references and nuances of close synonyms; employs structural precision, with sophisticated features that are readily, accurately, and appropriately controlled (such as complex modification and embedding in Indo-European languages); has discourse competence in a wide range of contexts and tasks, often matching a native speaker's strategic and organizational abilities and expectations. Occasional patterned errors occur in low frequency and highly-complex structures.

S-4 ADVANCED PROFESSIONAL PROFICIENCY (Base Level)

Able to use the language fluently and accurately on all levels normally pertinent to professional needs. An S-4's language usage and ability to function are fully successful. Organizes discourse well, employing functional rhetorical speech devices, native cultural references, and understanding. Language ability only rarely hinders him/her in performing any task requiring language; yet, an S-4 would seldom be perceived as a native. Speaks effortlessly and smoothly and is able to use the language with a high degree of effectiveness, reliability, and precision for all representational purposes within the range of personal and professional experience and scope of responsibilities. Can serve as an informal interpreter in a range of unpredictable circumstances. Can perform extensive, sophisticated language tasks, encompassing most matters of interest to well-educated native speakers, including tasks which do not bear directly on a professional specialty.

Examples: Can discuss in detail concepts which are fundamentally different from those of the target culture and make those concepts clear and accessible to the native speaker. Similarly, an S-4 can understand the details and ramifications of concepts that are culturally or conceptually different from his/her own. Can set the tone of interpersonal official, semi-official, and non-professional verbal exchanges with a representative range of native speakers (in a range of varied audiences, purposes, tasks, and settings). Can play an effective role among native speakers in such contexts as conferences, lectures, and debates on matters of disagreement. Can advocate a position at length, both formally and in chance encounters, using sophisticated verbal strategies. Can understand and reliably produce shifts of both subject matter and tone. Can understand native speakers of the standard and other major dialects in essentially any face-to-face interaction.

S-4+ ADVANCED PROFESSIONAL PROFICIENCY (Higher Level)

Speaking proficiency is regularly superior in all respects, usually equivalent to that of a well-educated, highly articulate native speaker. Language ability does not impede the performance of any language-use task. However, an S-4+ would not necessarily be perceived as culturally native.

Examples: An S-4+ organizes discourse well, employing functional rhetorical speech devices, native cultural references and understanding. Effectively applies a native speaker's social and circumstantial knowledge. However, cannot sustain that performance under all circumstances. While an S-4+ has a wide range and control of structure, an occasional non-native slip may occur. An S-4+ has a sophisticated control of vocabulary and phrasing that is rarely imprecise, yet there are occasional weaknesses in idioms, colloquialisms, pronunciation, cultural reference or there may be an occasional failure to interact in a totally native manner.

S-5 FUNCTIONALLY NATIVE PROFICIENCY

Speaking proficiency is functionally equivalent to that of a highly articulate well-educated native speaker and reflects the cultural standards of the country where the language is natively spoken. An S-5 uses the language with complete flexibility and intuition, so that speech on all levels is fully accepted by well-educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialisms, and pertinent cultural references. Pronunciation is typically consistent with that of well-educated native speakers of a non-stigmatized dialect.

INTERAGENCY LANGUAGE ROUNDTABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
LISTENING

L-0 NO PROFICIENCY

No practical understanding of the spoken language. Understanding is limited to occasional isolated words with essentially no ability to comprehend communication.

L-0+ MEMORIZED PROFICIENCY

Sufficient comprehension to understand a number of memorized utterances in areas of immediate needs. Slight increase in utterance length understood but requires frequent long pauses between understood phrases and repeated requests on the listener's part for repetition. Understands with reasonable accuracy only when this involves short memorized utterances or formulae. Utterances understood are relatively short in length. Misunderstandings arise due to ignoring or inaccurately hearing sounds or word endings (both inflectional and non-inflectional), distorting the original meaning. Can understand only with difficulty even persons such as teachers who are used to speaking with non-native speakers. Can understand best those statements where context strongly supports the utterance's meaning. Gets some main ideas.

L-1 ELEMENTARY PROFICIENCY (Base Level)

Sufficient comprehension to understand utterances about basic survival needs and minimum courtesy and travel requirements. In areas of immediate need or on very familiar topics, can understand simple questions and answers, simple statements and very simple face-to-face conversations in a standard dialect. These must often be delivered more clearly than normal at a rate slower than normal, with frequent repetitions or paraphrase (that is, by a native used to dealing with foreigners). Once learned, these sentences can be varied for similar level vocabulary and grammar and still be understood. In the majority of utterances, misunderstandings arise due to overlooked or misunderstood syntax and other grammatical clues. Comprehension vocabulary inadequate to understand anything but the most elementary needs. Strong interference from the candidate's native language occurs. Little precision in the information understood owing to tentative state of passive grammar and lack of vocabulary. Comprehension areas include basic needs such as: meals, lodging, transportation, time and simple directions (including both route instructions and orders from customs officials, policemen, etc.). Understands main ideas.

L-1+ ELEMENTARY PROFICIENCY (Higher Level)

Sufficient comprehension to understand short conversations about all survival needs and limited social demands. Developing flexibility evident in understanding into a range of circumstances beyond immediate survival needs. Shows spontaneity in understanding by speed, although consistency of understanding uneven. Limited vocabulary range necessitates repetition for understanding. Understands commoner time forms and most question forms, some word order patterns but miscommunication still occurs with more complex patterns. Cannot sustain understanding of coherent structures in longer utterances or in unfamiliar situations. Understanding of descriptions and the giving of precise information is limited. Aware of basic cohesive features, e.g., pronouns, verb inflections, but many are unreliably understood, especially if less immediate in reference. Understanding is largely limited to a series of short, discrete utterances. Still has to ask for utterances to be repeated. Some ability to understand the facts.

L-2 LIMITED WORKING PROFICIENCY (Base Level)

Sufficient comprehension to understand conversations on routine social demands and limited job requirements. Able to understand face-to-face speech in a standard dialect, delivered at a normal rate with some repetition and rewording, by a native speaker not used to dealing with foreigners, about everyday topics, common personal and family news, well-known current events, and routine office matters through descriptions and narration about current, past and future events; can follow essential points of discussion or speech at an elementary level on topics in his/her special professional field. Only understands occasional words and phrases of statements made in unfavorable conditions, for example through loudspeakers outdoors. Understands factual content. Native language causes less interference in listening comprehension. Able to understand the facts, i.e., the lines but not between or beyond the lines.

L-2+ LIMITED WORKING PROFICIENCY (Higher Level)

Sufficient comprehension to understand most routine social demands and most conversations on work requirements as well as some discussions on concrete topics

related to particular interests and special fields of competence. Often shows remarkable ability and ease of understanding, but under tension or pressure may break down. Candidate may display weakness or deficiency due to inadequate vocabulary base or less than secure knowledge of grammar and syntax. Normally understands general vocabulary with some hesitant understanding of everyday vocabulary still evident. Can sometimes detect emotional overtones. Some ability to understand between the lines (i.e., to grasp inferences).

L-3 GENERAL PROFESSIONAL PROFICIENCY (Base Level)

Able to understand the essentials of all speech in a standard dialect including technical discussions within a special field. Has effective understanding of face-to-face speech, delivered with normal clarity and speed in a standard dialect, on general topics and areas of special interest; understands hypothesizing and supported opinions. Has broad enough vocabulary that rarely has to ask for paraphrasing or explanation. Can follow accurately the essentials of conversations between educated native speakers, reasonably clear telephone calls, radio broadcasts, news stories similar to wire service reports, oral reports, some oral technical reports and public addresses on non-technical subjects; can understand without difficulty all forms of standard speech concerning a special professional field. Does not understand native speakers if they speak very quickly or use some slang or dialect. Can often detect emotional overtones. Can understand between the lines (i.e., grasp inferences)

L-3+ GENERAL PROFESSIONAL PROFICIENCY (Higher Level)

Comprehends most of the content and intent of a variety of forms and styles of speech pertinent to professional needs, as well as general topics and social conversation. Ability to comprehend many sociolinguistic and cultural references. However, may miss some subtleties and nuances. Increased ability to comprehend unusually complex structures in lengthy utterances and to comprehend many distinctions in language tailored for different audiences. Increased ability to understand native speakers talking quickly, using nonstandard dialect or slang; however, comprehension not complete. Some ability to understand "beyond the lines" in addition to strong ability to understand "between the lines."

L-4 ADVANCED PROFESSIONAL PROFICIENCY (Base Level)

Able to understand all forms and styles of speech pertinent to professional needs. Able to understand fully all speech with extensive and precise vocabulary, subtleties and nuances in all standard dialects on any subject relevant to professional needs within the range of his/her experience, including social conversations; all intelligible broadcasts and telephone calls; and many kinds of technical discussions and discourse. Understands language specifically tailored (including persuasion, representation, counseling, and negotiating) to different audiences. Able to understand the essentials of speech in some non-standard dialects. Has difficulty in understanding extreme dialect and slang, also in understanding speech in unfavorable conditions, for example through bad loudspeakers out doors. Understands "beyond the lines" all forms of the language directed to the general listener, (i.e., able to develop and analyze the argumentation presented).

L-4+ ADVANCED PROFESSIONAL PROFICIENCY (Higher Level)

Increased ability to understand extremely difficult and abstract speech as well as ability to understand all forms and styles of speech pertinent to professional needs, including social conversations. Increased ability to comprehend native speakers using extreme nonstandard dialects and slang as well as to understand speech in unfavorable conditions. Strong sensitivity to sociolinguistic and cultural references. Accuracy is close to that of the well-educated native listener but still not equivalent.

L-5 FUNCTIONALLY NATIVE PROFICIENCY

Comprehension equivalent to that of the well-educated native listener. Able to understand fully all forms and styles of speech intelligible to the well-educated native listener, including a number of regional and illiterate dialects, highly colloquial speech and conversations and discourse distorted by marked interference from other noise. Able to understand how natives think as they create discourse. Able to understand extremely difficult and abstract speech.

INTERAGENCY LANGUAGE ROUNDTABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
READING

Preface

In the following descriptions a standard set of text-types is associated with each level. The text-type is generally characterized in each descriptive statement.

The word "read," in the context of these proficiency descriptions, means that the person at a given skill level can thoroughly understand the communicative intent in the text-types described. In the usual case the reader could be expected to make a full representation, thorough summary, or translation of the text into English.

Other useful operations can be performed on written texts that do not require the ability to "read," as defined above. Examples of such tasks which persons of a given skill level may reasonably be expected to perform are provided, when appropriate, in the descriptions.

R-0 NO PROFICIENCY

No practical ability to read the language. Consistently misunderstands or cannot comprehend at all.

R-0+ MEMORIZED PROFICIENCY

Can recognize all the letters in the printed version of an alphabetic system and high-frequency elements of a syllabary or a character system. Able to read some or all of the following: numbers, isolated words and phrases, personal and place names, street signs, office and shop designations. The above often interpreted inaccurately. Unable to read connected prose.

R-1 ELEMENTARY PROFICIENCY (Base Level)

Sufficient comprehension to read very simple connected written material in a form equivalent to usual printing or typescript. Can read either representations of familiar formulaic verbal exchanges or simple language containing only the highest frequency structural patterns and vocabulary, including shared international vocabulary items and cognates (when appropriate). Able to read and understand known language elements that have been recombined in new ways to achieve different meanings at a similar level of simplicity. Texts may include simple narratives of routine behavior; highly predictable descriptions of persons, places or things; and explanations of geography and government such as those simplified for tourists. Some misunderstandings possible on simple texts. Can get some main ideas and locate prominent items of professional significance in more complex texts. Can identify general subject matter in some authentic texts.

R-1+ ELEMENTARY PROFICIENCY (Higher Level)

Sufficient comprehension to understand simple discourse in printed form for informative social purposes. Can read material such as announcements of public events, simple prose containing biographical information or narration of events, and straightforward newspaper headlines. Can guess at unfamiliar vocabulary if highly contextualized, but with difficulty in unfamiliar contexts. Can get some main ideas and locate routine information of professional significance in more complex texts. Can follow essential points of written discussion at an elementary level on topics in his/her special professional field.

In commonly taught languages, an R-1+ may not control the structure well. For example, basic grammatical relations are often misinterpreted, and temporal reference may rely primarily on lexical items as time indicators. Has some difficulty with the cohesive factors in discourse, such as matching pronouns with referents. May have to read materials several times for understanding.

R-2 LIMITED WORKING PROFICIENCY (Base Level)

Sufficient comprehension to read simple, authentic written material in a form equivalent to usual printing or typescript on subjects within a familiar context. Able to read with some misunderstandings straightforward, familiar, factual material, but in general insufficiently experienced with the language to draw inferences directly from the linguistic aspects of the text. Can locate and understand the main ideas and details in material written for the general reader. However, persons who have professional knowledge of a subject may be able to summarize or perform sorting and locating tasks with written texts that are well beyond their general proficiency level. The R-2 can read uncomplicated, but authentic prose on familiar subjects that are normally presented in a predictable sequence which aids the reader in understanding. Texts may include descriptions and narrations in contexts such as news items describing frequently occurring events, simple biographical information, social notices, formulaic business letters, and simple technical material written for the general reader. Generally the prose that can be read by an R-2 is predominantly in straightforward/high-frequency sentence patterns. The R-2 does not have a broad active vocabulary (that is, which he/she recognizes immediately on sight), but is able to use contextual and real-world cues to understand the text. Char-

acteristically, however, the R-2 is quite slow in performing such a process. Is typically able to answer factual questions about authentic texts of the types described above.

R-2+ LIMITED WORKING PROFICIENCY (Higher Level)

Sufficient comprehension to understand most factual material in non-technical prose as well as some discussions on concrete topics related to special professional interests. Is markedly more proficient at reading materials on a familiar topic. Is able to separate the main ideas and details from lesser ones and uses that distinction to advance understanding. The R-2+ is able to use linguistic context and real-world knowledge to make sensible guesses about unfamiliar material. Has a broad active reading vocabulary. The R-2+ is able to get the gist of main and subsidiary ideas in texts which could only be read thoroughly by persons with much higher proficiencies. Weaknesses include slowness, uncertainty, inability to discern nuance and/or intentionally disguised meaning.

R-3 GENERAL PROFESSIONAL PROFICIENCY (Base Level)

Able to read within a normal range of speed and with almost complete comprehension a variety of authentic prose material on unfamiliar subjects. Reading ability is not dependent on subject matter knowledge, although it is not expected that an R-3 can comprehend thoroughly subject matter which is highly dependent on cultural knowledge or which is outside his/her general experience and not accompanied by explanation. Texttypes include news stories similar to wire service reports or international news items in major periodicals, routine correspondence, general reports, and technical material in his/her professional field; all of these may include hypothesis, argumentation,and supported opinions. Misreading rare. Almost always able to interpret material correctly, relate ideas, and "read between the lines," (that is, understand the writers' implicit intents in texts of the above types). Can get the gist of more sophisticated texts, but may be unable to detect or understand subtlety and nuance. Rarely has to pause over or reread general vocabulary. However, may experience some difficulty with unusually complex structure and low frequency idioms.

R-3+ GENERAL PROFESSIONAL PROFICIENCY (Higher Level)

Can comprehend a variety of styles and forms pertinent to professional needs. Rarely misinterprets such texts or rarely experiences difficulty relating ideas or making inferences. Able to comprehend many sociolinguistic and cultural references. However, may miss some nuances and subtleties. Able to comprehend a considerable range of intentionally complex structures, low frequency idioms, and uncommon connotative intentions; however, accuracy is not complete. The S-3+ is typically able to read with facility, understand, and appreciate contemporary expository, technical, or literary texts which do not rely heavily on slang and unusual idioms.

R-4 ADVANCED PROFESSIONAL PROFICIENCY (Base Level)

Able to read fluently and accurately all styles and forms of the language pertinent to professional needs. The R-4's experience with the written language is extensive enough that he/she is able to relate inferences in the text to real-world knowledge and understand almost all sociclinguistic and cultural references. Able to "read beyond the lines" (that is, to understand the full ramifications of texts as they are situated in the wider cultural, political, or social environment). Able to read and understand the intent of writers' employment of nuance and subtlety. An R-4 can discern relationships among sophisticated written materials in the context of broad experience. Can follow unpredictable turns of thought readily in, for example, editorial, conjectural, and literary texts in any subject matter area directed to the general reader. Can read essentially all materials in his/her special field, including official and professional documents and correspondence. Recognizes all professionally relevant vocabulary known to the educated non-professional native, although may have some difficulty with slang. Can read reasonably legible handwriting without difficulty. Accuracy is often nearly that of a well-educated native reader.

R-4+ ADVANCED PROFESSIONAL PROFICIENCY (Higher Level)

Nearly native ability to read and understand extremely difficult or abstract prose, a very wide variety of vocabulary, idioms, colloquialisms, and slang. Strong sensitivity to and understanding of sociolinguistic and cultural references. Little difficulty in reading less than fully legible handwriting. Broad ability to "read beyond the lines" (that is, to understand the full ramifications of texts as they are situated in the wider cultural, political, or social environment) is nearly that of a well-read or well-educated native reader. Accuracy is close to that of the well-educated native reader, but not equivalent.

R-5 FUNCTIONALLY NATIVE PROFICIENCY

Reading proficiency is functionally equivalent to that of the well-educated native reader. Can read extremely difficult and abstract prose; for example, general legal and technical as well as highly colloquial writings. Able to read literary texts, typically including contemporary avant-garde prose, poetry, and theatrical writing. Can read classical/archaic forms of literature with the same degree of facility as the well-educated, but non-specialist native. Reads and understands a wide variety of vocabulary and idioms, colloquialisms, slang, and pertinent cultural references. With varying degrees of difficulty, can read all kinds of handwritten documents. Accuracy of comprehension is equivalent to that of a well-educated native reader.

INTERAGENCY LANGUAGE ROUNDTABLE
LANGUAGE SKILL LEVEL DESCRIPTIONS
WRITING

W-0 NO PROFICIENCY

No functional writing ability.

W-0+ MEMORIZED PROFICIENCY

Writes using memorized material and set expressions. Can produce symbols in an alphabetic or syllabic writing system or 50 of the most common characters. Can write numbers and dates, own name, nationality, address, etc., such as on a hotel registration form. Otherwise, ability to write is limited to simple lists of common items such as a few short sentences. Spelling and even representation of symbols (letters, syllables, characters) may be incorrect.

W-1 ELEMENTARY PROFICIENCY (Base Level)

Has sufficient control of the writing system to meet limited practical needs. Can create by writing statements and questions on topics very familiar to him/her within the scope of his/her very limited language experience. Writing vocabulary is inadequate to express anything but elementary needs; writes in simple sentences making continual errors in spelling, grammar and punctuation but writing can be read and understood by a native reader used to dealing with foreigners attempting to write his/her language. Writing tends to be a loose collection of sentences (or fragments) on a given topic and provides little evidence of conscious organization. While topics which are "very familiar" and elementary needs vary considerably from individual to individual, any person at this level should be able to write simple phone messages, excuses, notes to service people and simple notes to friends. (800-1000 characters controlled.)

W-1+ ELEMENTARY PROFICIENCY (Higher Level)

Sufficient control of writing system to meet most survival needs and limited social demands. Can create sentences and short paragraphs related to most survival needs (food, lodging, transportation, immediate surroundings and situations) and limited social demands. Can express fairly accurate present and future time. Can produce some past verb forms but not always accurately or with correct usage. Can relate personal history, discuss topics such as daily life, preferences and very familiar material. Shows good control of elementary vocabulary and some control of basic syntactic patterns but major errors still occur when expressing more complex thoughts. Dictionary usage may still yield incorrect vocabulary or forms, although the W-1+ can use a dictionary to advantage to express simple ideas. Generally cannot use basic cohesive elements of discourse to advantage (such as relative constructions, object pronouns, connectors, etc.). Can take notes in some detail on familiar topics, and respond to personal questions using elementary vocabulary and common structures. Can write simple letters, summaries of biographical data and work experience with fair accuracy. Writing, though faulty, is comprehensible to native speakers used to dealing with foreigners.

W-2 LIMITED WORKING PROFICIENCY (Base Level)

Able to write routine social correspondence and prepare documentary materials required for most limited work requirements. Has writing vocabulary sufficient to express himself/herself simply with some circumlocutions. Can write simply about a very limited number of current events or daily situations. Still makes common errors in spelling and punctuation but shows some control of the most common formats and punctuation conventions. Good control of morphology of language (in inflected languages) and of the most frequently used syntactic structures. Elementary constructions are usually handled quite accurately and writing is understandable to a native reader not used to reading the writing of foreigners. Uses a limited number of cohesive devices.

W-2+ LIMITED WORKING PROFICIENCY (Higher Level)

Shows ability to write with some precision and in some detail about most common topics. Can write about concrete topics relating to particular interests and special fields of competence. Often shows surprising fluency and ease of expression but under time constraints and pressure language may be inaccurate and/or incomprehensible. Generally strong in either grammar or vocabulary but not in both. Weaknesses or unevenness in one of the foregoing or in spelling result in occasional miscommunication. Areas of weakness range from simple constructions such as plurals, articles, prepositions and negatives to more complex structures such as tense usage, passive constructions, word order and relative clauses. Normally controls general vocabulary with some misuse of everyday vocabulary evident. Shows a limited ability to use circumlocutions. Uses dictionary to advantage to supply unknown words. Can take fairly accurate notes on material presented orally and handle with fair accuracy most social correspondence. Writing is understandable to native speakers not used to dealing with foreigners' attempts to write the language, though style is still obviously foreign.

W-3 GENERAL PROFESSIONAL PROFICIENCY (Base Level)

Able to use the language effectively in most formal and informal written exchanges on practical, social and professional topics. Can write reports, summaries, short library research papers on current events, on particular areas of interest or on special fields with reasonable ease. Control of structure, spelling and general vocabulary is adequate to convey his/her message accurately but style may be obviously foreign. Errors virtually never interfere with comprehension and rarely disturb the native reader. Punctuation generally controlled. Employs a full range of structures. Control of grammar good with only sporadic errors in basic structures, occasional errors in the most complex frequent structures and somewhat more frequent errors in low frequency complex structures. Consistent control of compound and complex sentences. Relationship of ideas is consistently clear.

W-3+ GENERAL PROFESSIONAL PROFICIENCY (Higher Level)

Able to write the language in a few prose styles pertinent to professional/educational needs. Not always able to tailor language to suit audience. Weaknesses may lie in poor control of low frequency complex structures, vocabulary or the ability to express subtleties and nuances. May be able to write on some topics pertinent to professional/educational needs. Organization may suffer due to lack of variety in organizational patterns or in variety of cohesive devices.

W-4 ADVANCED PROFESSIONAL PROFICIENCY (Base Level)

Able to write the language precisely and accurately in a variety of prose styles pertinent to professional/educational needs. Errors of grammar are rare including those in low frequency complex structures. Consistently able to tailor language to suit audience and able to express subtleties and nuances. Expository prose is clearly, consistently and explicitly organized. The writer employs a variety of organizational patterns, uses a wide variety of cohesive devices such as ellipsis and parallelisms, and subordinates in a variety of ways. Able to write on all topics normally pertinent to professional/educational needs and on social issues of a general nature. Writing adequate to express all his/her experiences.

W-4+ ADVANCED PROFESSIONAL PROFICIENCY (Higher Level)

Able to write the language precisely and accurately in a wide variety of prose styles pertinent to professional/educational needs. May have some ability to edit but not in the full range of styles. Has some flexibility within a style and shows some evidence of a use of stylistic devices.

W-5 FUNCTIONALLY NATIVE PROFICIENCY

Has writing proficiency equal to that of a well-educated native. Without non-native errors of structure, spelling, style or vocabulary can write and edit both formal and informal correspondence, official reports and documents, and professional/educational articles including writing for special purposes which might include legal, technical, educational, literary and colloquial writing. In addition to being clear, explicit and informative, the writing and the ideas are also imaginative. The writer employs a very wide range of stylistic devices.

21 November 1983

These descriptions were approved by the Interagency Language Roundtable, consisting of the following agencies.

Department of Defense
Department of State
Central Intelligence Agency
National Security Agency
Department of the Interior
National Endowment for the Humanities
National Institutes of Health
National Science Foundation
Department of Agriculture
Drug Enforcement Administration
Federal Bureau of Investigation
ACTION/Peace Corps
Agency for International Development
Office of Personnel Management
Immigration and Naturalization Service
Department of Education
US Customs Service
US Information Agency
Library of Congress

TOEFL® Research Reports currently available . . .

Report 1. *The Performance of Native Speakers of English on the Test of English as a Foreign Language.* John L. D. Clark. November 1977.

Report 2. *An Evaluation of Alternative Item Formats for Testing English as a Foreign Language.* Lewis W. Pike. June 1979.

Report 3. *The Performance of Non-Native Speakers of English on TOEFL and Verbal Aptitude Tests.* Paul J. Angelis, Spencer S. Swinton, and William R. Cowell. October 1979.

Report 4. *An Exploration of Speaking Proficiency Measures in the TOEFL Context.* John L. D. Clark and Spencer S. Swinton. October 1979.

Report 5. *The Relationship between Scores on the Graduate Management Admission Test and the Test of English as a Foreign Language.* Donald E. Powers. December 1980.

Report 6. *Factor Analysis of the Test of English as a Foreign Language for Several Language Groups.* Donald E. Powers and Spencer S. Swinton. December 1980.

Report 7. *The Test of Spoken English as a Measure of Communicative Ability in English-Medium Instructional Settings.* John L. D. Clark and Spencer S. Swinton. December 1980.

Report 8. *Effects of Item Disclosure on TOEFL Performance.* Gordon A. Hale, Paul J. Angelis, and Lawrence A. Thibodeau. December 1980.

Report 9. *Item Performance Across Native Language Groups on the Test of English as a Foreign Language.* Donald L. Alderman and Paul W. Holland. August 1981.

Report 10. *Language Proficiency as a Moderator Variable in Testing Academic Aptitude.* Donald L. Alderman. November 1981.

Report 11. *A Comparative Analysis of TOEFL Examinee Characteristics, 1977-1979.* Kenneth M. Wilson. July 1982.

Report 12. *GMAT and GRE Aptitude Test Performance in Relation to Primary Language and Scores on TOEFL.* Kenneth M. Wilson. July 1982.

Report 13. *The Test of Spoken English as a Measure of Communicative Ability in the Health Professions: Validation and Standard Setting.* Donald E. Powers and Charles W. Stansfield. January 1983.

Report 14. *A Manual for Assessing Language Growth in Instructional Settings.* Spencer S. Swinton. February 1983.

Report 15. *Survey of Academic Writing Tasks Required of Graduate and Undergraduate Foreign Students.* Brent Bridgeman and Sybil Carlson. September 1983.

Report 16. *Summaries of Studies Involving the Test of English as a Foreign Language, 1963-1982.* Gordon A. Hale, Charles W. Stansfield, and Richard P. Duran. February 1984.

Report 17. *TOEFL from a Communicative Viewpoint on Language Proficiency: A Working Paper.* Richard P. Duran, Michael Canale, Joyce Penfield, Charles W. Stansfield, and Judith E. Liskin-Gasparro. February 1985.

If you wish additional information about TOEFL research or would like to be placed on the mailing list to automatically receive order forms for newly published reports, write to:

TOEFL Program Office
CN 6155
Princeton, NJ 08541-6155
USA

BEST COPY AVAILABLE

5790609 • RP1575 • L76572 • Printed in U.S.A.